# CAREERS

**GRANTS** Success rate for UK government funding is lower for women **p.525**

**SALARIES** Faculty pay up more at US public universities than at private **p.525**

**NATUREJOBS** For the latest career listings and advice www.naturejobs.com

**RESEARCH TOOLS**

# Jump off the page

*Researchers are learning to embrace online lab notebooks, but not without growing pains.*

**BY AMANDA MASCARELLI**

Two years into his PhD, Carl Boettiger needed a better way to organize his data and synthesize his ideas. Fishing around online, he stumbled across chemist Cameron Neylon's open electronic lab notebook. Boettiger, who was studying mathematical ecology, liked what he saw. Neylon, now advocacy director at the Public Library of Science in San Francisco, California, had pulled back the curtain on the steps and thought processes behind his protocols and research. His data collection, protocols and results were linked together and available online, making the concepts easy to reference and explore.

Inspired, Boettiger created his own electronic notebook, reporting online about his day-to-day research in a publicly available wiki that

is followed by the open-science community. Viewers can find the notebook through the wiki's RSS feed and links on social media as well as in Google searches, and can post comments. He soon started to receive suggestions about, and valuable feedback on, his research and methods from other scientists — mostly followers of open science from outside his field — and some even led to collaborations.

Four years on, Boettiger, now a theoretical ecologist at the University of California, Santa Cruz, is a leader in the use of the open notebook. He is convinced of its value, and he is not alone — the idea is steadily gaining traction in some (but not all) scientific circles.

Whether on paper or in digital form, lab notebooks are meant to document exactly what, when and why experiments were done (see 'Record of achievement') and usually

contain much more information than will ever be published in an academic paper. They can be used as evidence for securing patents, to settle legal issues or to pass a project from one researcher to another. Industry labs almost always require their researchers to maintain such records, as do many academic principal investigators, and until very recently the information would be kept securely under wraps in the lab until it was published.

Open electronic notebooks are a radical departure from this ethos. Data and methods are no longer cloistered in books or tucked away on private hard drives. Instead, they can be shared online for all to see. Some scientists might shudder at the thought of anyone but lab mates and close collaborators knowing the detailed logic and steps behind their research projects before publication. But open ▶

science is becoming more widely accepted as technologies change and as younger generations of researchers discover alternative tools and approaches.

## EMBRACE THE OPENNESS

Boettiger readily admits that having his entire scientific process exposed online at each step of the way carries a risk. But "you have to take risks to be successful", he points out. "The idea that there's a risk-free way of getting your science out there and understood and engaging collaborators is a myth." He posts updates to his work throughout the day, as if scribbling notes in a conventional notebook, but publishes synthesized analyses and summaries several times a week or month. Boettiger still has some concerns about being scooped (see *Nature* **493**, 711; 2013), but says that those reservations are mostly offset by the many benefits of making his science open and accessible — such as providing another way to attract collaborators and to gain recognition in his field.

The gradual acceptance of open science in fields such as chemistry, mathematics, neuroscience and ecology is highlighting the information-management challenges that scientists face. "Now that science has become more complicated and not all of the details and data fit in the papers we publish, we're at a loss," says Boettiger. "How do we communicate exactly what we're doing? How do we keep science replicable?"

Scientists collect, store and analyse data in different ways and with a growing array of tools. That makes it difficult to compare one researcher's results with another's, and even harder to know how to reproduce their results, says Carly Strasser, a data-curation and open-science specialist at the California Digital Library in Oakland. "We still have paper notebooks, we still have Post-it notes, we still have phone calls that we don't transcribe and a ton of e-mails going back and forth," she says. "There are just a lot of moving parts. Scientists don't ever really learn how to capture the process."

The problem is mainly down to a lack of training. Undergraduates turn in detailed notebooks for relatively simple experiments, but as the complexity of their work increases when they become graduate students, they may struggle to document these larger experiments. Many labs do not have formalized notebook-writing conventions in place. Strasser says that her approach as a PhD student in marine biology was makeshift — cutting and pasting computer printouts into a paper notebook, storing data in "chaotic heaps" on her computer and using the comment feature in Excel to describe results.



*"There are just a lot of moving parts. Scientists don't ever really learn how to capture the process."*
Carly Strasser

## GO WITH THE FLOW

As researchers move beyond simply documenting their daily tasks, they are turning to electronic lab notebooks that have myriad bells and whistles and offer 'workflow' capabilities. These software packages aim to capture the entire scientific process — from study design to data analysis and visualization — in close to real time, and enable colleagues to review and replicate the work. Users can share data and methods as they develop, or choose to wait until publication.

A piece of software called R, for example, is widely used in the ecology community. With it, researchers can weave together text, programming code and data analysis into a narrative. The steps leading to a published paper no longer need to be impenetrable to other researchers — anyone can access any step along the way and reproduce it, or take the work in other directions. "These tools allow me to explore alternative ideas in a structured and documented way without disrupting my existing work," says Karthik Ram, a computational ecologist at the University of California, Berkeley.

Between 2009 and 2011, Ram was a post-doc at the University of California, Santa Cruz, studying the impacts of climate change on large mammals in Yellowstone National Park in Wyoming. He needed to incorporate decades of data from multiple sources — on the natural history of herbivores and on how migratory behaviour has changed over time, for example — and link it all to long-term climate data and changes in snowpack and vegetation. Because he uses existing data and models to answer questions, he often needs to understand intermediate steps along the way, such as the statistical methods used. Ram found it time-consuming and difficult to tease apart other researchers' data sets so that he could build on them. Workflow platforms such as R and IPython (see 'Taming the workflow') are now making this type of work much more manageable, he says. Other platforms include Projects (developed by Digital Science, a sister company of Nature Publishing Group).

Once a workflow system is populated with the nitty-gritty of the project, scientists can take advantage of Internet tools, such as wikis and social media, to post and share their work. Jean-Claude Bradley, an organic chemist at Drexel University in Philadelphia, Pennsylvania, has several projects in his lab that rely on wikis. Any raw data a project obtains, including pictures, videos and spectroscopy results, are quickly shared among team members, usually within a day. Lab members also use the wikis to share their findings. One wiki page represents one page in a lab notebook, with sections, objectives, methods and other components.

All these tools help to create what open-notebook aficionados call a living paper, which contains all components of the research, from e-mails to conference chatter

---

## RECORD OF ACHIEVEMENT

### *A glossary of electronic notebooks*

There are many ways to build a twenty-first-century electronic notebook — and to make work accessible to others.

● **Electronic notebook (open or closed)**
A digital version of the conventional paper lab notebook in which the entire scientific process can be captured. Unlike conventional notebooks, which are difficult for others to access, electronic notebooks make it easier for researchers to organize, manage and share the many components of their work.

● **Workflow software** Platforms such as R and IPython integrate all the pieces of research into a single system. They capture data collection and analyses in near real time and these can then be shared with colleagues for review or collaboration simply by clicking a button.

● **Wiki** Web tools that serve as platforms for electronic notebooks. Wikipedia, the online encyclopaedia, is one example of a wiki. A wiki can be made accessible and modifiable by others. Many researchers who practice open science use wikis to record their

progress and then share it with others. For example, a site called OpenWetWare allows researchers to build a notebook wiki.

● **Living paper** A term coined by researchers to describe the concept of archiving the entire scientific process, from phone calls and Twitter conversations to methods and analysis, in a way that is openly accessible. It allows other researchers to view the steps leading up to the final products and even to build on or derive ideas from the original work. **A.M.**

## TAMING THE WORKFLOW
### *The open-access platform IPython*

In 2001, as a final-year PhD student in theoretical physics, Fernando Perez started on a side project. For his research on quantum field theory, he needed to bring together his hotchpotch of computer code and data-analysis tools. Using a programming language called Python, he created IPython (ipython.org), an open-source, integrated platform that allowed him to type code, run his analyses, plot and visualize his data and add rich graphics within a single system.

"When I started working on IPython I told myself and my adviser that this was just going to be an afternoon hack and that I would get back to 'real work' very soon," he says. Some 13 years later, Perez is a computational scientist at the Brain Imaging Center at the University of California, Berkeley, and he develops IPython for computational science, publication and education across domain disciplines as his full-time job.

In 2011, he, along with collaborator Brian Granger at California Polytechnic State University in San Luis Obispo and their colleagues, added in a web-enabled notebook, which has been rapidly adopted by computational scientists working in the fields of biology, physics and neuroscience. It functions like a word processor, with normal text and formatting, but also enables users to insert programming language and rich graphics and data analyses, and easily go back and forth between them. "It's like having a very powerful calculator in the middle of your word processor that can do anything programming language can do," says Perez.

Researchers have even begun to publish papers directly from IPython, says Perez. The University of California, Berkeley, offers courses in IPython, and Harvard University and the Massachusetts Institute of Technology in Cambridge, and Columbia University in New York, among others, have adopted it.

The National Center for Ecological Analysis and Synthesis at the University of California, Santa Barbara, runs workshops on collaborative synthesis and data-sharing for graduate students and postdocs. And short, on-site workshops by a worldwide volunteer group called Software Carpentry offer hands-on training in Python, R, GitHub and other data-synthesis programs at various locations. **A.M.**

on Twitter, and links them together in an openly accessible, easily updatable, digital workflow. "A living paper is alive in that it gets updated and it reflects the ongoing process," says Boettiger. It allows scientists to document their entire scientific process, which they — and whomever they choose to share it with — can add to and derive new ideas from, says Matthew Jones, an expert in informatics at the National Center for Ecological Analysis and Synthesis at the University of California, Santa Barbara.

### PRESERVED FOR POSTERITY
All these records must be stored, and archiving not only data and computations, but also intellectual discussions, presents challenges. Boettiger links his Twitter feed to his open notebook, and generates a running online tab of his reading and notes through the Mendeley reference manager. "Our literature has expanded so rapidly that other people are often our best sources of being able to figure out what to read," he says. Peter Andras, a computer scientist at Newcastle University, UK, also logs his reading through Mendeley and encourages his students to do the same.

But web-based resources often decline in popularity and if that happens, the record can languish unseen, or potentially even be lost. FriendFeed, for example, a hybrid between Twitter and Facebook, had a healthy researcher following in 2010 but now has substantially fewer users. Phillip Lord, a bioinformatician and a collaborator of Andras, says that for security he archives digital files through both archive.org, a general web repository, and an initiative at the British Library in London. Archiving services act as networks of libraries, storing data in multiple places to ensure that if one copy were lost or destroyed, it could be retrieved elsewhere.

Such challenges make it difficult to maintain electronic and open notebooks — but they are unlikely to stop their increasing adoption. "We're moving away from the science that we can document with a pen and paper to the science that we do in six different venues on instruments and with GPSs and laptops and paper and pen," says Strasser. "How do we capture that? It's super hard." Even so, the advantages are hard to ignore. "The days of paper lab books," Lord says, "are well past their best." ∎

*Amanda Mascarelli is a freelance writer based in Denver, Colorado.*

# Gender differences

Women are less successful than men in getting grants from Research Councils UK (RCUK), according to an 11 March report from the umbrella body for the UK government's research councils. RCUK's first such analysis compared the annual success rate during 2010–13 of grant applications from women and men in eight age groups. Across all age and grant categories, women averaged a 25% success rate, compared with men's 29%. Iain Cameron, head of research careers and diversity at RCUK, says that the reason for women's lower success is not clear. One council is launching unconscious-bias training for peer reviewers, and RCUK may do a follow-up analysis. "The point of publishing these data is to expose them," Cameron says.

# Swiss frozen out

Early-career researchers in Switzerland will have limited access to the Erasmus Plus student-exchange programme and other European research funds for the next academic year. Switzerland lost its status as an Erasmus Plus programme country after it voted to impose quotas on immigrants from European Union (EU) countries that restrict the free movement of people from Croatia, the newest EU member. The change means that Swiss nationals who wish to study in the EU will no longer be eligible for Erasmus grants and that students from eligible countries will not be able to obtain an Erasmus grant to study in Switzerland. Erasmus funds around 7,000 exchanges a year between Switzerland and other countries.

# Public boost

Pay rises were slightly higher for faculty members at US public institutions (2.2%) than at private institutions (2.0%) for the first time since 2009, finds the College and University Professional Association for Human Resources (CUPA-HR) in Knoxville, Tennessee. The change suggests an overall strengthening of state economies, says Andy Brantley, CUPA-HR president. CUPA-HR's poll of 792 institutions and 180,000 faculty members also showed that compensation for tenure-track and tenured faculty members across both university categories rose by an average of 2.1% this academic year, the same increase as the year before.