Jocelyn Bell Burnell discovered pulsars as a PhD student at the University of Cambridge, UK.

BREAKTHROUGH PRIZE

Pulsar discoverer wins $3–million prize

Jocelyn Bell Burnell to use cash to promote diversity in science.

BY ZEEYA MERALI

Fifty years after discovering pulsars — compact rotating stars that emit beams of radiation — astrophysicist Jocelyn Bell Burnell has been awarded one of the most lucrative prizes in science: a US$3-million Breakthrough prize. Thought by many to have been snubbed for a Nobel prize for the discovery, Bell Burnell, 75, has been recognized by the Breakthrough committee with a special award in fundamental physics both for her scientific achievements and for her "inspiring leadership" over the past five decades.

"I cannot think of a more deserving scientist to win this prize," says Chiara Mingarelli, an astrophysicist at the Flatiron Institute in New York City. "In addition to being both a pioneer and a giant in the field, Bell Burnell is the highest calibre role model — a champion for women in science."

Bell Burnell, now at the University of Oxford, UK, and the University of Dundee, UK, was "totally surprised" to learn about the prize, which was announced on 6 September. "For once in my life, I was speechless," she says. She is already in discussions with the national physics institutes in the United Kingdom and Ireland about using the prize money to create PhD studentships for people from under-represented groups in science. "Diversity is very important," says Bell Burnell. "This also recognizes that I did my most important work as a student."

The Breakthrough prizes were launched in 2012 and are funded by entrepreneurs including Google co-founder Sergey Brin. Awarded in fundamental physics, life sciences and mathematics, they are usually handed out in December, based on selections made after an open nomination process. But the selection committee can decide to make special awards. Previous special awards have been given to Stephen Hawking and to the Laser Interferometer Gravitational-Wave Observatory collaboration for the discovery of gravitational waves.

Pulsars are dense stars, consisting mostly of neutrons, that rotate at a precise rate, emitting radiation as they spin. In 1967, Bell Burnell, then a PhD student at the University of Cambridge, UK, under astronomer Antony Hewish, was analysing hundreds of metres of chart paper containing data collected by a radio telescope in Cambridge when she noticed some mysterious recurring smudges. She was able to characterize these as signs of radio pulses emanating from a spinning star: the pulsar. "The discovery is a testament to her curiosity, her determination and her creativity," says Mingarelli.

In 1974, Hewish shared the Nobel Prize in Physics with fellow radio astronomer Martin Ryle, for pioneering research in astrophysics. Hewish was cited for his "decisive role in the discovery of pulsars" — while Bell Burnell was overlooked. Bell Burnell herself has previously stated that she does not mind the oversight because she understands that Nobel prizes are not usually awarded to research students.

Her discovery still rings out half a century on, notes Mingarelli. For instance, earlier this year, NASA scientists demonstrated the feasibility of using pulsars to navigate, with their SEXTANT experiment on the International Space Station. The idea is that future robotic spacecraft could use the clockwork-like arrival times of X-rays emitted from pulsars to triangulate their locations. ∎

INFORMATION

Google enters data ecosystem

Dataset Search could be especially helpful to cross-disciplinary researchers.

BY DAVIDE CASTELVECCHI

Google has unveiled a search engine to help researchers locate online data that are freely available for use. The company launched the service on 5 September, saying that it is aimed at "scientists, data journalists, data geeks, or anyone else".

Dataset Search, now available alongside Google's other specialized search engines, such as those for news and images — as well as Google Scholar and Google Books — locates files and databases according to how their owners have classified them. It does not read the content of the files themselves in the way search engines do for web pages.

Experts say that it fills a gap and could contribute significantly to the success of the open-data movement, which aims to make data freely available for use and re-use.

Government agencies, scientific publishers, research institutions and individual researchers maintain thousands of open-data repositories worldwide, containing millions of data sets.

But researchers who want to know what types of data are available, or who hope to locate data they know already exist, often have to rely on word of mouth, says Natasha Noy, a computer scientist at Google AI in Mountain View, California.

▶

DAVID HARTLEY/SHUTTERSTOCK

▶ This problem is especially serious for early-career researchers who are not already "plugged" into a network of professional connections, Noy says. It's also a downside for those who do cross-disciplinary research — for example, an epidemiologist who needs access to climate data that could be relevant to the spread of a virus.

CLASSIFIED SEARCH

Noy and her Google colleague Dan Brickley first described a strategy for solving that problem in a blogpost in January 2017 (see go.nature.com/2oounre).

Typical search engines work in two main stages. The first is to index the available pages by continuously trawling the Internet. The second is to rank those indexed pages, so that when a user enters search terms, the engine can provide results in order of relevance.

To aid search engines in indexing existing data sets, Noy and Brickley wrote, those who own the data sets should 'tag' them, using a standardized vocabulary called Schema.org, an initiative founded by Google and three other search-engine giants (Microsoft, Yahoo and Yandex), and which Brickley manages. The Google team also developed a special algorithm for ranking data sets in search results.

Given Google's dominance in web searching, news that the company was moving into the data ecosystem quickly prompted major players to fall in line and standardize their metadata, says Mark Hahnel, chief executive of the data-sharing company Figshare in London. (Figshare is operated by the Holtzbrinck Publishing Group, which also has a majority share in *Nature*'s publisher.)

"By November, all the universities we're working for had their stuff marked up," Hahnel says. "I think this is a game changer for open data in the academia."

Funding agencies sometimes mandate that research data be made available, and they are going to reach their ultimate goals only if the information is effectively recoverable, he says. "It legitimizes what the funders have been trying to do."

AGENCY PARTNERSHIPS

An early supporter of Google's experiment was the US National Oceanic and Atmospheric Administration (NOAA). The agency's remit ranges from fisheries to the Sun's corona, and its archives contain nearly 70,000 data sets — including ship logs from the 1800s. The trove adds up to more than 35 petabytes, comparable to the content of 35,000 typical hard drives.

Google's tool will help NOAA to meet its open-data mission, says NOAA's chief data officer Edward Kearns in Asheville, North Carolina. "We want to explore new ways to make those data available to others," Kearns adds.

For Dataset Search to work, having the data owners' collaboration was a crucial step. Although the system might become more sophisticated in the future, Google currently has no plans to actually read the data or analyse them, as it does with web pages or images. "A search tool like this one is only as good as the metadata that data publishers are willing to provide," Noy says.

Like Google Scholar, Dataset Search currently offers no access for automated querying and no application programming interface (API) — although the company says that it might add that functionality in the future.

Noy says that as researchers begin to use Dataset Search, Google will watch how they interact with it and use that information to improve the search results. The company has no current plans to monetize the service, she says.

As Dataset Search evolves, it might also become integrated with Google Scholar, so that search results on a particular study could link to relevant data sets. ∎

CORRECTION

The Editorial 'What is Life?' (*Nature* **561**, 6; 2018) misnamed Trinity College Dublin.