ESSAY

The Harvard computers

The first mass data crunchers were people, not machines. **Sue Nelson** looks at the discoveries and legacy of the remarkable women of Harvard's Observatory.

A photograph taken at the Harvard Observatory in Cambridge, Massachusetts, *circa* 1890, features eight women in what looks like a Victorian-style sitting room. They wear long skirts, have upswept hair and are surrounded by flowered wallpaper and mahogany tables. At first glance they seem to be sampler stitching or reading. In fact these 'human computers' are analysing photographs of the heavens, cataloguing stars.

When cameras were first attached to telescopes, with the ability to capture the image of thousands of stars on a single photographic plate, people were needed to trawl through these new data. Observatories hired 'computers' — a term used for human processors since the early 1700s — to do the painstakingly repetitive work of measuring the brightness, position and colours of these stars.

From the 1880s until the 1940s, the Harvard College Observatory amassed half a million photographic glass plates, weighing around 300 tonnes and holding images of tens of millions of stars. A team of women trawled through these photos with nothing more than magnifying glasses — often for little pay and with no scientific training.


HARVARD COLL. OBSERV.

Williamina Fleming stands in the centre of the Harvard computers as Edward Pickering looks on.

Despite these unpromising conditions, the 'Harvard computers', who worked from the end of the nineteenth century to the mid-1920s, made tremendous contributions to astronomy. They determined how to calculate the vast distances from Earth to the stars, and developed star classification systems that are still used today. From photos taken of the northern and southern skies, from observatories in Cambridge, New Zealand and Peru, they produced an astronomical gold mine of data.

These women were proof that 'people power', even from those with no formal training, is capable of great things. It is a trend that continues today: volunteers are recruited from the general population and taught to spot objects of interest to astronomers, from the tracks of interstellar dust left in a spacecraft's collector, to the direction of spin of a spiral galaxy. With Harvard now working to digitize its photographic plates, the same pictures of stars scrutinized by the Harvard computers may soon be available to many more, equally curious, eyes.

Working with the repetitive and often indistinct photographs collected at places such as the Harvard Observatory required patience, attention to detail and stamina. Most of the plates are negatives; stars appear as fine grey or black spots against a clear background. There are also several thousand spectral plates, in which starlight has been split by a prism before being captured. These look like nothing more than smudged pencil marks a few millimetres wide; under a magnifying glass the smudge turns into a barcode, revealing information about the chemical composition and temperature of the stars.

Patience personified

In 1901, William Elkin, the director of Yale Observatory, expressed a view typical of the time as to who was best suited for this work. "I am thoroughly in favour of employing women as measurers and computers," he said. "Not only are women available at smaller salaries than are men, but for routine work they have important advantages. Men are more likely to grow impatient after the novelty of the work has worn off and would be harder to retain for that reason."

Edward Pickering, the Harvard College Observatory director in 1877–1919, famously said that the computing work at his observatory was so easy that even his "Scotch maid" could do it. This was Williamina Fleming, a schoolteacher from Dundee who had emigrated to America with her husband in 1878. A year later, abandoned and pregnant, she secured a job as Pickering's maid and housekeeper. She was soon working for him at the observatory part-time as one of his first computers.

Pickering's apparently disparaging remark about his maid belies the fact that he spotted and nourished the untapped potential in many intelligent women who worked for him. Fleming was obviously bright and Pickering recognized this; by 1881, at the age of 24, she was appointed a full-time staff member of Harvard. Seven years later, she assumed responsibility for the increasing number of photographic plates, editing publications from the observatory and hiring new computers. During her time at Harvard, Fleming examined thousands of spectra and catalogued more than 10,000 stars.

Fleming helped Pickering to devise his hydrogen-based stellar classification system, which ranked stars according to the strength of a hydrogen spectral line — A for the strongest, then B and so on. She also played a crucial part in the discovery of the spectral peculiarities of white dwarfs. Fleming was appointed Harvard's Curator of Astronomical Photographs in 1899 — the first woman so appointed — and eight years later she became the first female American citizen elected to the British Royal Astronomical Society. During her career it is estimated that she examined around 200,000 photographic plates.

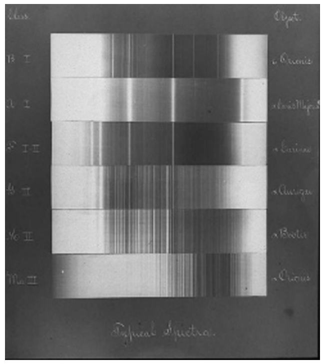
Pickering's 'harem', as they have been called, could have earned more per hour doing menial work in the local mill town. But the observatory appealed to intelligent and educated women, including researchers and graduates from the new all-female colleges, who were keen to find patterns in the data and draw interesting conclusions from them. Pickering allowed the women to do their own research in their spare time and their names were often cited as co-authors in scientific papers. He encouraged them to give talks, and to be recognized as astronomers in their own right.

Annie Jump Cannon, for example, measured and recorded the colours of 300,000 stars, classifying them into spectral groups at a rate of up to 300 an hour. Her true achievement came when she was tasked with finding a more meaningful way of arranging the star categories in Pickering's hydrogen-based system. Cannon developed a reordering and simplification that ranked stars in order from the bluest and hottest to the coolest, red ones. The sequence — O, B, A, F, G, K, M (remembered by the mnemonic Oh Be A Fine Girl, Kiss Me) — remains in use today.

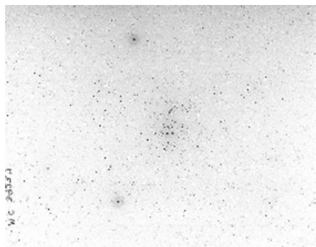
Henrietta Swan Leavitt, a college graduate, joined Pickering as a research assistant. She studied Cepheid variable stars — whose light brightens and dims over periods ranging from hours to years — and found that these periods were related to the bodies' intrinsic brightness in a predictable way. This relationship, once calibrated, allowed astronomers to determine such a star's distance, based on its true brightness as compared to its apparent brightness as seen from Earth. This 'Cepheid distance scale' became Edwin Hubble's 'yardstick to the Universe', eventually allowing him to discover that the Universe was expanding. It is still crucially important in distance measurements today.

Leavitt died in 1921 of cancer, four years before a letter was sent notifying her that she had been nominated for a Nobel prize — an honour that had to be withdrawn because it could not be awarded posthumously. In his 2005 book *The Discoveries*, physicist Alan Lightman nominates the Cepheid distance scale as one of the most important breakthroughs in twentieth-century science. To celebrate the centennial of Leavitt's pioneering work, the Harvard College Observatory is holding a symposium in her honour this November.

More advanced cameras and digital photography removed the need for most photographic plates in the 1980s. But the sheer number of plates collected at Harvard means that even now there are stars that have yet to be analysed. Because of this untapped potential, and the historical significance of the computers' work, the DASCH (Digital Access to a Sky Century at Harvard) project is hoping to scan all the half-million plates over the next four years; the team has so far scanned more than 3,000 plates and is trying to raise US$4 million to complete the work.



HARVARD COLL. OBSERV.



HARVARD COLL. OBSERV.

Spectra (top) pulled from star plates (above), are only millimetres wide.

Pattern recognition

The primary motivation of this work is not historical preservation, but to dig deeper into the data. Requests have already been filed by researchers for access to the DASCH's digitized collection. One Harvard graduate student, Sumin Tang, working with DASCH principal investigator Josh Grindlay, recently found something unusual after studying just 500 of the digitized plates: a star that brightened by nearly a factor of two over 20 years and then levelled off, retaining its brightness for 60 years. This new type of variable star will be observed further using a telescope in Arizona in the years to come.

"There remain some areas in which the human eye is better than a modern computer."

The sheer magnitude of data expected for the entire DASCH project (around 1,500 terabytes) lends itself well to analysis by today's high-speed computers. Yet there remain some areas in which the human eye is better than a modern computer, particularly in the realm of pattern recognition. Classifying stars or galaxies, like classifying species, is still something most easily done by people with knowledge and a knack.

In some cases, the best results come from the use of many eyes — even if they are untrained. NASA's 2004 Stardust mission, for example, gathered dust particle samples on an aerogel collector during a rendezvous with a comet. One of the greatest challenges then lay in finding the tracks of tiny interstellar particles — rare pieces of dust that came from distant stars — among the more common tracks left by comet particles. These tracks, at a millionth of a metre across, were a devil to find on the 1,000-square-centimetre collector.

In 2006, the Planetary Society and the University of California, Berkeley, launched the Stardust@Home project. Just as the Harvard computers were trained to analyse stellar spectra, members of the public were trained, via online tutorials, to scan photos of the gel on their computer screens and identify possible tracks. No computer program exists that can do this as well as the human eye.

In the first phase of the project, 23,000 volunteer 'dusters' searched nearly 40 million images, flagging any photos of possible interest for trained scientists. Without public help, it would have taken the team at least 20 years to locate tracks. With their help, it took months. Phase two is currently under way, with dusters searching for particles in photos of a higher magnification.

Inspired by the Stardust project, Galaxy Zoo went online in 2007. This project used volunteers to classify spiral and elliptical galaxies from images taken by the Sloan Digital Sky Survey. Within six months of operation, volunteers had identified more than 500 overlapping galaxies — when a galaxy appears behind or in front of another from the point of view of an observer on Earth — of which only 20 were previously known to astronomers. Recently, a unique object containing the hot gas of a normal galaxy but without any stars was discovered through the project by a Dutch primary school teacher; a paper describing the find is in the works.

Modern astronomy could not be done without supercomputers crunching epic quantities of data. It is nonetheless worth remembering the value of the human mind in spotting complex patterns or following a hunch — as proved by the persistent, repetitive and inspiring work of the pioneering women at Harvard. ■

Sue Nelson is a writer and broadcaster living in Hertfordshire, UK.
e-mail: sue.nelson@zen.co.uk

See Editorial, page 1.