AI reads text from ancient Herculaneum scroll for the first time

Machine-learning technique reveals Greek words in CT scans of rolled-up papyrus.

⏱ 6 min. czytania

Charred scrolls from Herculaneum can't be opened easily, but X-ray scanning can reveal their contents.Credit: UK Photo

A 21-year-old computer-science student has won a global contest to read the first text inside a carbonized scroll from the ancient Roman city of Herculaneum, which had been unreadable since a volcanic eruption in ad 79 — the same one that buried nearby Pompeii. The breakthrough could open up hundreds of texts

from the only intact library to survive from Greco-Roman antiquity.

Luke Farritor, who is at the University of Nebraska–Lincoln, developed a ==machine-learning algorithm== that has detected Greek letters on several lines of the rolled-up papyrus, including πορφυρας (*porphyra*s), meaning 'purple'. Farritor used subtle, small-scale differences in surface texture to train his neural network and highlight the ink.

"When I saw the first image, I was shocked," says Federica Nicolardi, a papyrologist at the University of Naples in Italy and a member of the academic committee that reviewed Farritor's findings. "It was such a dream," she says. Now, "I can actually see something from the inside of a scroll."

Hundreds of scrolls were buried by Mount Vesuvius in October ad 79, when the eruption left Herculaneum under 20 metres of volcanic ash. Early attempts to open the papyri created a mess of fragments, and scholars feared the remainder could never be unrolled or read. "These are such crazy objects. They're all crumpled and crushed," says Nicolardi.

The Vesuvius Challenge offers a series of awards, leading to a main prize of US$700,000 for reading four or more passages from a rolled-up scroll. On 12 October, the organizers announced that Farritor has won the 'first

letters' prize of $40,000 for reading more than 10 characters in a 4-square-centimetre area of papyrus. Youssef Nader, a graduate student at the Free University of Berlin, is awarded $10,000 for coming second.

Luxury library

To finally see letters and words inside a scroll is "extremely exciting", says Thea Sommerschield, a historian of ancient Greece and Rome at Ca' Foscari University of Venice, Italy. The scrolls were discovered in the eighteenth century, when workmen came across the remains of a luxury villa that might have belonged to the family of Julius Caesar's father-in-law. Deciphering the papyri, Sommerschield says, could "revolutionize our knowledge of ancient history and literature". Most classical texts known today are the result of repeated copying by scribes over centuries. By contrast, the Herculaneum library contains works not known from any other sources, direct from the authors.

Until now, researchers were able to study only opened fragments. A few Latin works have been identified, but most of these contain Greek texts relating to the Epicurean school of philosophy. There are parts of *On Nature*, written by Epicurus himself, and works by a little-known philosopher named Philodemus on topics such as vices, music, rhetoric and death. It has been suggested that the library might once have been his working collection. But

more than 600 scrolls — most held in the National Library in Naples, with a handful in the United Kingdom and France — remain intact and unopened. And more papyri could still be found on lower floors of the villa, which have yet to be excavated.

Brent Seales, a computer scientist who has helped set up the Vesuvius Challenge, and his team spent years developing methods to "virtually unwrap" the vanishingly thin layers using X-ray computed tomography (CT) scans, and to visualize them as a series of flat images. In 2016 Seales, who is at the University of Kentucky in Lexington, he reported[1] using the technique to read a charred scroll from En-Gedi in Israel, revealing sections of the Book of Leviticus — part of the Jewish Torah and the Christian Old Testament — written in the third or fourth century ad. But the ink on the En-Gedi scroll contains metal, so it glows brightly on the CT scans. The ink on the older Herculaneum scrolls is carbon-based, essentially charcoal and water, with the same density in scans as the papyrus it sits on, so it doesn't show up at all.

Seales realized that even with no difference in brightness, CT scans might capture tiny differences in texture that can distinguish areas of papyrus coated with ink. To prove it, he trained an artificial neural network to read letters in X-ray images of opened Herculaneum

[handwritten annotation:] how?

fragments. Then, in 2019, he carried two intact scrolls from the Institut de France in Paris to the Diamond Light Source, a synchrotron X-ray facility near Oxford, UK, to scan them at the highest resolution yet (4–8 micrometres per 3D image element, or voxel).

[handwritten annotation: What is this]

Reading intact scrolls was still a huge task, however, so the team released all of its scans and code to the public and launched the Vesuvius Challenge. "We all agreed we would rather get to the reading of what's inside sooner, than try to hoard everything," says Seales.

Around 1,500 teams were soon discussing and collaborating through the gamer chat platform Discord. The prizes were designed in phases, and as each milestone is reached, the winning code is released for everyone to build on. Farritor, who had always been interested in history and taught himself Latin as a child, got involved early on.

In parallel, Seales' team worked on the virtual unwrapping, releasing images of the flattened pieces for the contestants to analyse. A key moment came in late June, when one competitor pointed out that on some images, ink was occasionally visible to the naked eye, as a subtle texture that was soon dubbed 'crackle'. Farritor immediately focused on the crackle, looking for further hints of letters.

One evening in August, he was at a party when he received an alert that a fresh segment had been released, with particularly prominent crackle. Connecting through his phone, he ran his algorithm on the new image. Walking home an hour later, he pulled out his phone and saw five letters on the screen. "I was jumping up and down," he says. "Oh my goodness, this is actually going to work." From there, it took just days to refine the model and identify the ten letters required for the prize.
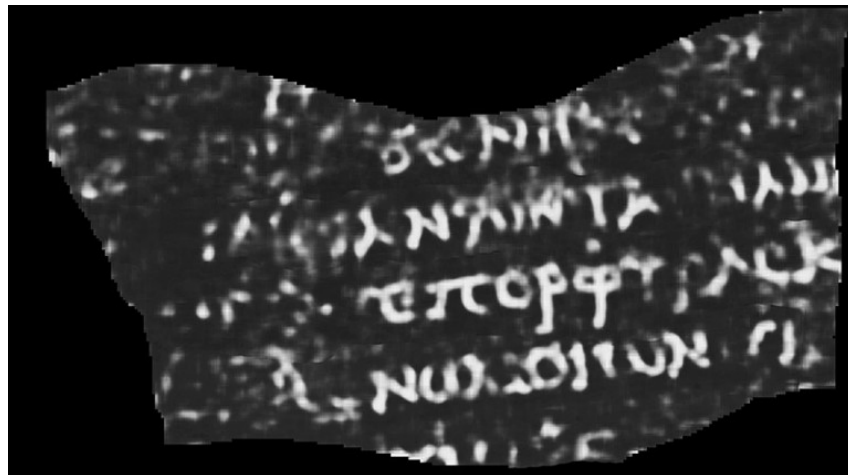
The first word deciphered in the Vesuvius Challenge is the Greek πορφύρας, meaning 'purple'.Credit: UK Photo

Papyrologists are excited, too. The word "purple" has not yet been read in the opened Herculaneum scrolls. Purple dye was highly sought-after in ancient Rome and was made from the glands of sea snails, so the term could refer to purple colour, robes, the rank of people who could afford the dye or even the molluscs. But more important than the individual word is reading anything at all, says Nicolardi. The

advance "gives us potentially the possibility to recover the text of a whole scroll", including the title and author, so that works can be identified and dated.

Seeing the invisible

Yannis Assael, a staff research scientist at Google DeepMind in London, describes the Vesuvius Challenge as "unique and inspirational". But it is part of a broader shift, he notes, in which ==artificial intelligence== (AI) is increasingly aiding the study of ancient texts. Last year, for example, Assael and Sommerschield [released an AI tool called Ithaca](#), designed to help scholars glean the date and origins of unidentified ancient Greek inscriptions, and make suggestions for text to fill any gaps[2]. It now receives hundreds of ==queries== per week, and similar efforts are being applied to languages from Korean to Akkadian, which was used in ancient Mesopotamia.

Seales hopes ==machine learning== will open up what he calls the "invisible library". This refers to texts that are physically present, but no one can see, including parchment used in medieval book bindings; palimpsests, in which later writing obscures a layer beneath; and cartonnage, in which scraps of old papyrus were used to make ancient Egyptian mummy cases and masks.

For now, however, all eyes are on the Vesuvius Challenge. The deadline for the grand prize is 31 December, and Seales describes the mood as "unbridled optimism". Farritor, for one, has already run his models on other segments of the scroll and is seeing many more characters appear.

Updates & Corrections

- **Correction 13 October 2023**: An earlier version of this article did not include the full name and affiliation of Brent Seales.