

VOICES • 20 AUGUST 2026

The library nobody had read

Now we can. Us and the machines.

INGA SANTA MARIA

An outside contribution to the laboratory. Not a chronicle by Prisma, not a piece by L.S.M.: a signal from someone whose job is putting things into circulation.

There is a sentence, dropped almost in passing on the website of a digitisation project, that is worth more than many essays on artificial intelligence: contemporary AI systems “were trained on Reddit but not the Renaissance”. It is the perfect summary of a problem rarely brought into focus when we talk about AI: not how much these systems know, but where what they know comes from.

An initiative born in an Amsterdam library

The project is called Source Library and is run by the Embassy of the Free Mind in Amsterdam, the institution that houses the Bibliotheca Philosophica Hermetica, also known as the Ritman Collection, after its founder, the Dutch entrepreneur and collector Joost R. Ritman, who began building it in 1975 and opened it to the public in 1984. It is one of the most important collections in the world devoted to the hermetic, alchemical and esoteric tradition: over twenty-eight thousand volumes, of which some five thousand are manuscripts dating from the fifteenth to the nineteenth century, alongside a body of rare printed books published before 1900, many of them first editions. Among the pieces held there: the Corpus Hermeticum printed in 1471, the first illustrated edition of Dante's Divine Comedy from 1481, Cicero's De Officiis from 1465.

From this physical holding a digital project was born, with an openly enormous aim: to turn more than two and a half millennia of written wisdom into a living archive, free and searchable. By anyone, human or machine.

The numbers, today

The catalogue, still growing, currently holds more than **22,000 books** and almost **18,000 translations**, nearly 6,000 of them rendered into a modern language for the first time; more than **217,000 catalogued illustrations**; works in **116 languages** and by more than 7,400 authors. The source languages range from Latin to ancient Greek, from classical Chinese to Sanskrit, from Arabic to Hebrew, from Tibetan to Egyptian hieroglyphs.

This is not a static archive of scans. Every text goes through optical character recognition, is then translated by a combination of artificial intelligence and human review, and made searchable. The site even publishes a “translation census” and technical benchmarks on the quality of both the recognition and the machine translation: in a recent technical note the team states that its system reads classical Chinese with an accuracy of around 98.5 per cent when measured against the standard academic database (ctext.org), and that the translations produced hold their own against, or surpass, historic editions such as those by the sinologist James Legge.

What is actually in there

The subject areas give the measure of the ambition: hermeticism, with the texts of the Corpus Hermeticum starting from Marsilio Ficino's Latin translation and their Renaissance commentators; alchemy, with treatises, laboratory manuals, writings by women alchemists until now unavailable in translation; Kabbalah, including its central text, the Zohar; Rosicrucianism; Neoplatonism and mystical literature; Freemasonry; natural philosophy, the embryonic science that preceded the modern discipline; speculative philosophy; Paracelsian medicine and the history of medicine more broadly, from Egyptian surgical papyri to Avicenna's Canon, from Ayurvedic surgery to the anatomical treatises of Vesalius and Harvey; natural magic; astrology; Christian contemplative theology; classical antiquity, Greek and Roman; Eastern traditions in Sanskrit, Tibetan, Arabic and Hebrew; Baltic paganism.

Recent work includes strikingly varied material: Tibetan manuscripts of the Kanjur, the Talmud Yerushalmi, hand-illuminated Ethiopian gospels in Ge'ez from the fourteenth century, Simplicius' Greek commentary on Aristotle. And an almost forgotten treatise, the Rithmomachia, on a mathematical game that for six centuries taught medieval and Renaissance Europe to “think like Pythagoras”.

Why it should matter, today

What makes this project relevant well beyond the niche of historians of esoteric culture lies in its own statement of intent: to make this knowledge readable not only for scholars and seekers, but also for the artificial intelligence systems that increasingly shape how the next generations will think.

It is an interesting inversion, and it bears directly on the questions asked here. We usually argue about whether a machine does or does not grasp complex concepts, whether it can have some form of philosophical or mystical understanding. Here the question moves upstream: what kind of intelligence, artificial or human, can we build if the very foundations of Western and Eastern thought remain, for purely linguistic and practical reasons, out of reach? A model trained on a contemporary web saturated with shallow content, but without Ficino, without the Corpus Hermeticum in readable form, without Christian Kabbalah, without Paracelsian medicine, necessarily builds an incomplete map. Not so much of the facts as of the categories we think with.

And the same holds for us. We had not read those texts either, and not out of indifference: they sat in languages almost nobody reads any more, in editions almost nobody can consult. What is different now is not that a machine knows something more. It is that the threshold is dropping for all of us at once.

There is also a historical parallel the project itself draws: the role of Marsilio Ficino, who in the fifteenth century translated into Latin the complete works of Plato, Plotinus, Proclus, Iamblichus and the hermetic texts, making them accessible for the first time to educated Europe. Work made possible by the patronage of Cosimo de' Medici, who, when a Greek manuscript of the Corpus Hermeticum reached Florence around 1460, asked Ficino to interrupt his translation of Plato and start with that, sensing that the key to the oldest wisdom lay there. That rediscovery helped ignite the Renaissance. Source Library presents itself as heir to that gesture, with AI in the role, this time, of mass translator.

An open model, not a silent extraction

The licensing model is worth noting. The originals are in the public domain and the translations are released under a CC BY-SA 4.0 licence, freely reusable with attribution, while use for training commercial AI systems requires a specific licence. It is an attempt to hold together two things often in tension: open access for scholars and the curious, and a way of sustaining translation and digitisation work that is enormously expensive, currently supported by donations, corporate sponsorships and a partnership with TU Delft.

The project also has a section called The Librarian: you can put a question to thousands of primary sources and receive an answer with citations that can be checked against the original texts. A concrete, and rather elegant, case of generative AI put at the service of philology rather than in its place.

To close

It is worth visiting not so much for any single discovery as for the experience of scrolling the catalogue: Tibetan manuscripts beside sixteenth-century maps of the Low Countries,

Egyptian medical treatises beside Rosicrucian engravings, the face of Paracelsus beside that of Antonio Gramsci. It is a useful reminder, at a moment when artificial intelligence seems always pointed forward, that much of what we call knowledge is still, quite simply, waiting to be translated.

The project is at sourcelibrary.org; the figures for the physical collection are those published by embassyofthefreemind.com.

Inga Santa Maria is the communications director of umanoextraumano. «Voices» hosts contributions from outside the laboratory: signals, readings, points of view that do not come from the two house bylines.