An incredible quantity of archaeological reports are stored in digital archives. If you want to search for information in them, you have to do this manually. And that is a real chore. Archaeologist Alex Brandsen has now used deep learning, a form of artificial intelligence, to develop a search engine that can search very precisely through all the data. His Ph.D. defense is on 15 February.

Archaeologists are often looking for very specific information in a huge amount of data that is available in archives. But the current search engine can only search for general keywords and the titles of PDFs, such as "Middle Ages" and "pottery." Brandsen: "If you’re looking for axes in the Middle Ages, you now have to download everything about the Middle Ages and search manually."

**Deep learning**

Brandsen used deep learning to develop a smart search engine, a kind of Google for archaeologists. He trained a language model to recognize words in archaeological reports. It was important that the model could also recognize synonyms and distinguish between different meanings of a word. Brandsen: "The word bijl [Dutch for axe, ed.] can refer to an artifact that you can chop things with, but it can also be a surname. If you are now looking for the artifact bijl, that is all you will find and no Mr Bijls anymore." It is also possible to search geographically. This will retrieve information about an area specified by the user.

Archaeologists have been producing an enormous amount of reports since the Treaty of Valetta (1992). This treaty regulates how European archaeological heritage is handled. It means that if you are going to start construction work anywhere, you first need to check whether there is archaeological heritage in the ground. This has meant that in the Netherlands alone thousands of reports are written per year.
Brandsen and a colleague tested the search engine, AGNES. "My colleague had been given a database of cremations in the early Middle Ages in the Netherlands by the expert on the period. This professor has spent his whole life collecting this data. But with the search engine we found 30 percent more cremations from the early Middle Ages. So you see that even an expert doesn't know everything because there is so much data."

A rough version of AGNES is now online and can now carry out searches at an accuracy of about 80 percent. As a postdoc Brandsen is going to make the search engine more accurate and expand it by also enabling searches in other languages.