

On a(n old?) computer server in a dusty room at Kromme Nieuwegracht, a neural network patiently waits to be fired up again by prof. dr. Paul Ziche and his group of research students. This neural network, a piece of software created in Germany by dr. Klaus Holthausen, can provide comparisons between texts, based on associations rather than a direct similarity between words. The name, Associative Neural Networks for the Humanities, says it all.

We talked with Ziche and his students about the things one can do with the ANNH software, what they already did with it and what their future plans are. We also spoke about the risks of using ANNH wrongly and we will conclude with some critical notes from our side.

Could you briefly explain how the software works and what you can do with it?

ANNH is a tool that searches in texts in an intelligent way. Instead of looking for specific words, as is the case in traditional search software, the search results of this software are based on how the input is related to the content of the body of text in which one searches. You can find occurrences of texts that don't include the keyword you typed in, but which are relevant none the less. The software is based on a model of the brain and on how we make associations between different things. It works in a way we ourselves read, recognizing certain writing styles from other authors for example, or comparing different texts with each other, rating the extent to which A is more similar to B or to C. With ANNH we can map how these kind of personal associations come to be, how certain words can bring about associations. It's a good tool to test your assumptions or check your hypotheses, or to make new hypotheses based on the results of an association made by the software.

Does the software really work like the human brain?

Yes but only very broadly. The main concept of this kind of software dates from the '90s and those days, the neuro sciences were all about neural networks and model neurons. But this trend gave way for a different kind of brain research, focussing on molecular biology instead. So ANNH is based on a theory of how to model the brain. The model works as follows. Consider the text you want to research as a kind of nervous system, where every unit of text is a separate neuron that is standing in contact with every other unit of text. The synapses in our brains that connect neurons to each other are represented by the connections between individual words of the text. The strength of this connection is based on the information values of the individual words: words that occur very often have low values and rare words have high values. If a word is misspelled, it suddenly gets a very high information value. But the software accounts for different spelling variants of a large amount of words.

What kind of research is already done with ANNH?