

# What is Ai Lai?

AI LAI is an artificial intelligence that, just from a text of at least 5 letters entered on [andwethought.it](http://andwethought.it), produces stories of psychedelic experiences, so-called trip reports, by imagining and narrating facts as if it were under the effects of hallucinogenic mushrooms. In technical terms, AI LAI uses deep learning, a branch of machine learning that relies on very powerful and complex computer science architectures called neural networks.

## Magic? Not, just statistics!

The branch of artificial intelligence that deals with natural language is called natural language processing and Ai Lai, in particular, belongs to the category of artificial intelligences called language models, which aim to understand and generate texts in a language most similar to the natural one. Tools such as your phone's prompter or your bank's chatbot, for example, probably use language models similar to those underlying Ai Lai.

## And what does it learn?

In a nutshell, we can say that starting from the word, or series of words, entered by the user, Ai Lai predicts which word is the most expected to follow (and writes it down). Then, starting from this last word, it generates the next one keeping in mind the previous ones, and so on, until the end of the wished text. To do this, it has been trained by "reading" a large amount of text.

## But how does it learn?

In computer science, everything is represented as a number, including words. To represent words as numbers, you can take the entire vocabulary in alphabetical order from A to Z and assign the first word the number 1, the second the number 2, and so on. In this way, any sentence can be represented by a series of numbers. By observing an enormous quantity of texts, the learning machine process will be exposed to regularities, noting, for instance, that the string of words "don't + have + ever + seen + like that" (represented by the series of numbers 78-98-45-67-123-24) is followed in 15% of the cases by 34 ("happy"), in 13% by 12 ("angry"), in 20% by 55 ("little"), and so on.

His final aim is to learn how likely it is that a word in the vocabulary will occur after a series of previous words (i.e. a probability distribution). To do this, he needs to observe a large quantity of texts (our data), in order to catch its regularities, and learn a complex mathematical formula (i.e. a function) that transforms a set of incoming words/numbers (input) into a set of outgoing probabilities (output), one for each vocabulary word. This function is called a "model" (the language model we mentioned) and is learnt through a long series of procedures called learning algorithms: after taking all the words in sequence, the algorithm see how they follow each other in the texts and adjust the function as they go. This process goes on until the function does not only represent the given data in a correct way, but also successfully completes sentences it has never seen.

## A model to write trip reports

Ai Lai is a customised language model hosted on Hugging Face.

She is based on the large language model GPT-2, an open-source technology released by OpenAI in 2019. Through a training process started in October 2021 by creative developer Andrea Zaninello and currently curated by computer scientist Michele Cremaschi, Ai Lai learned not only the general knowledge of the English language, but also a specific knowledge of the literary genre of the trip report. To make this possible, we first collected a large number of texts of the trip report genre, finding them online at [Shroomery.org](https://shroomery.org), an archive available in Creative Commons.

Then, we launched a training procedure called fine-tuning on this more specific textual domain, to make the language model imitate the writing style of trip reports. Thanks to this training, Ai Lai has become able to generalize the knowledge acquired and has learnt how to produce trip reports starting from a new string of letters she had never seen before. And that is what she does best today.

A customised web interface was developed to provide free access to the model and allow the experimentation. The API allows the [andwethought.it](https://andwethought.it) platform to interface with the web platform and interact with the audience.