

A **large language model (LLM)** is a computerized [language model](#) consisting of an [artificial neural network](#) with many parameters (tens of millions to billions), trained on large quantities of unlabeled text using [self-supervised learning](#) or [semi-supervised learning](#).^[1] LLMs emerged around 2018 and perform well at a wide variety of tasks. This has shifted the focus of [natural language processing](#) research away from the previous paradigm of training specialized [supervised](#) models for specific tasks.^[2]

Though the term *large language model* has no formal definition, it often refers to [deep learning](#) models with millions or even billions of parameters, that have been "pre-trained" on a large corpus. LLMs are general purpose models which excel at a wide range of tasks, as opposed to being trained for one specific task (such as [sentiment analysis](#), [named entity recognition](#), or [mathematical reasoning](#)).^{[2][3]} The skill with which they accomplish tasks, and the range of tasks at which they are capable, seems to be a function of the amount of resources (data, parameter-size, computing power) devoted to them, in a way that is not dependent on additional breakthroughs in design.^[4]

Though trained on simple tasks along the lines of predicting the next word in a sentence, neural language models with sufficient training and parameter counts are found to capture much of the syntax and semantics of human language. In addition, large language models demonstrate considerable general knowledge about the world, and are able to "memorize" a great quantity of facts during training.^[2]