**Понимание естественного языка** (**NLU**) или **интерпретация естественного языка** (**NLI**

**Понимание естественного языка** (**NLU**) или **интерпретация естественного языка** (**NLI**)[[1]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-1) — подраздел [обработки естественного языка](https://ru.wikipedia.org/wiki/%D0%9E%D0%B1%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%BA%D0%B0_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0) в [искусственном интеллекте](https://ru.wikipedia.org/wiki/%D0%98%D1%81%D0%BA%D1%83%D1%81%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D0%BB%D0%BB%D0%B5%D0%BA%D1%82), который имеет дело с машинным [пониманием текста](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D1%82%D0%B5%D0%BA%D1%81%D1%82%D0%B0). Понимание естественного языка является [AI-полной задачей](https://ru.wikipedia.org/wiki/AI-%D0%BF%D0%BE%D0%BB%D0%BD%D0%B0%D1%8F_%D0%B7%D0%B0%D0%B4%D0%B0%D1%87%D0%B0)[[2]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-2).

Существует значительный коммерческий интерес к этой области из-за её применения к автоматизированному рассуждению[[3]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-3), [машинному переводу](https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D1%8B%D0%B9_%D0%BF%D0%B5%D1%80%D0%B5%D0%B2%D0%BE%D0%B4)[[4]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-4), [ответам на вопросы](https://ru.wikipedia.org/wiki/%D0%92%D0%BE%D0%BF%D1%80%D0%BE%D1%81%D0%BD%D0%BE-%D0%BE%D1%82%D0%B2%D0%B5%D1%82%D0%BD%D0%B0%D1%8F_%D1%81%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0)[[5]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-5), сбору новостей, [классификации текста](https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D1%8F_%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D0%BE%D0%B2), [голосовой активации](https://ru.wikipedia.org/wiki/%D0%93%D0%BE%D0%BB%D0%BE%D1%81%D0%BE%D0%B2%D0%BE%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D1%80%D1%84%D0%B5%D0%B9%D1%81), архивированию и крупномасштабному [анализу контента](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BD%D1%82%D0%B5%D0%BD%D1%82-%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7).

**История**

Программа STUDENT, написанная в 1964 году Дэниелом Бобровым для его докторской диссертации в [Массачусетском технологическом институте](https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%81%D1%81%D0%B0%D1%87%D1%83%D1%81%D0%B5%D1%82%D1%81%D0%BA%D0%B8%D0%B9_%D1%82%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B8%D0%BD%D1%81%D1%82%D0%B8%D1%82%D1%83%D1%82), является одной из первых известных попыток компьютерного понимания естественного языка[[6]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-6)[[7]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-7)[[8]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-8)[[9]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-9)[[10]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-10). Через восемь лет после того, как [Джон Маккарти](https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D0%BA%D0%BA%D0%B0%D1%80%D1%82%D0%B8,_%D0%94%D0%B6%D0%BE%D0%BD) ввел термин «[искусственный интеллект](https://ru.wikipedia.org/wiki/%D0%98%D1%81%D0%BA%D1%83%D1%81%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D0%BB%D0%BB%D0%B5%D0%BA%D1%82)», в диссертации Боброва (озаглавленной «*Ввод данных на естественном языке для компьютерной системы решения задач»*) было показано, как компьютер может понимать простой ввод данных на естественном языке для решения задач по алгебре.

Год спустя, в 1965 году, [Джозеф Вейценбаум](https://ru.wikipedia.org/wiki/%D0%92%D0%B5%D0%B9%D1%86%D0%B5%D0%BD%D0%B1%D0%B0%D1%83%D0%BC,_%D0%94%D0%B6%D0%BE%D0%B7%D0%B5%D1%84) из Массачусетского технологического института написал [ELIZA](https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%B8%D0%B7%D0%B0_(%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B0)), интерактивную программу, которая вела диалог на английском языке на любую тему, самой популярной из которых была психотерапия. ELIZA работала путем простого синтаксического анализа и замены ключевых слов на готовые фразы, а Вейценбаум обошел проблему предоставления программе [базы данных](https://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%B7%D0%B0_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85) реальных знаний или богатого [словарного запаса](https://ru.wikipedia.org/wiki/%D0%9B%D0%B5%D0%BA%D1%81%D0%B8%D0%BA%D0%B0). Тем не менее ELIZA приобрела удивительную популярность как игрушечный проект и может рассматриваться как очень ранний предшественник современных коммерческих систем, таких как те, которые используются [Ask.com](https://ru.wikipedia.org/wiki/Ask.com)[[11]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-11).

В 1969 году [Роджер Шанк](https://ru.wikipedia.org/w/index.php?title=%D0%A0%D0%BE%D0%B4%D0%B6%D0%B5%D1%80_%D0%A8%D0%B0%D0%BD%D0%BA&action=edit&redlink=1) из [Стэнфордского университета](https://ru.wikipedia.org/wiki/%D0%A1%D1%82%D1%8D%D0%BD%D1%84%D0%BE%D1%80%D0%B4%D1%81%D0%BA%D0%B8%D0%B9_%D1%83%D0%BD%D0%B8%D0%B2%D0%B5%D1%80%D1%81%D0%B8%D1%82%D0%B5%D1%82" \o "Стэнфордский университет) представил концептуальную теорию зависимости для понимания естественного языка[[12]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-12). Эта модель, на которую частично повлияла работа Сиднея Лэмба, широко использовалась студентами Шанка в [Йельском университете](https://ru.wikipedia.org/wiki/%D0%99%D0%B5%D0%BB%D1%8C%D1%81%D0%BA%D0%B8%D0%B9_%D1%83%D0%BD%D0%B8%D0%B2%D0%B5%D1%80%D1%81%D0%B8%D1%82%D0%B5%D1%82), такими как Роберт Виленски, Венди Ленерт и Джанет Колоднер.

В 1970 году Уильям А. Вудс представил [сеть расширенных переходов](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D1%88%D0%B8%D1%80%D0%B5%D0%BD%D0%BD%D0%B0%D1%8F_%D1%81%D0%B5%D1%82%D1%8C_%D0%BF%D0%B5%D1%80%D0%B5%D1%85%D0%BE%D0%B4%D0%BE%D0%B2) (ATN) для представления ввода на естественном языке[[13]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-13). Вместо *правил структуры* фраз ATN использовали эквивалентный набор [автоматов с конечным состоянием](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BD%D0%B5%D1%87%D0%BD%D1%8B%D0%B9_%D0%B0%D0%B2%D1%82%D0%BE%D0%BC%D0%B0%D1%82), которые вызывались рекурсивно. ATN и их более общий формат, называемый «обобщенными ATN», продолжали использоваться в течение ряда лет.

В 1971 году [Терри Виноград](https://ru.wikipedia.org/wiki/%D0%92%D0%B8%D0%BD%D0%BE%D0%B3%D1%80%D0%B0%D0%B4,_%D0%A2%D0%B5%D1%80%D1%80%D0%B8) закончил писать [SHRDLU](https://ru.wikipedia.org/wiki/SHRDLU) для своей докторской диссертации в Массачусетском технологическом институте. SHRDLU мог понимать простые английские предложения в ограниченном мире детских блоков, чтобы направлять роботизированную руку для перемещения предметов. Успешная демонстрация SHRDLU придала значительный импульс дальнейшим исследованиям в этой области[[14]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-14)[[15]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-15). Виноград продолжал оказывать большое влияние в этой области с публикацией своей книги «*Язык как когнитивный процесс»*[[16]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-16). В Стэнфорде Виноград позже будет консультировать [Ларри Пейджа](https://ru.wikipedia.org/wiki/%D0%9F%D0%B5%D0%B9%D0%B4%D0%B6,_%D0%9B%D0%B0%D1%80%D1%80%D0%B8), сооснователя [Google](https://ru.wikipedia.org/wiki/Google_(%D0%BA%D0%BE%D0%BC%D0%BF%D0%B0%D0%BD%D0%B8%D1%8F)).

В 1970-х и 1980-х годах группа обработки естественного языка в SRI International продолжала исследования и разработки в этой области. На основе этого исследования был предпринят ряд коммерческих проектов, *например*, в 1982 году Гэри Хендрикс основал [Symantec Corporation](https://ru.wikipedia.org/wiki/NortonLifeLock), первоначально как компанию для разработки интерфейса на естественном языке для запросов к базе данных на персональных компьютерах. Однако с появлением [графических пользовательских интерфейсов](https://ru.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D1%84%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D1%80%D1%84%D0%B5%D0%B9%D1%81_%D0%BF%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8F), управляемых мышью, Symantec изменила направление. Ряд других коммерческих проектов был начат примерно в то же время, *например*, Ларри Р. Харрис в Корпорации искусственного интеллекта и Роджер Шанк и его ученики в Cognitive Systems Corp[[17]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-17)[[18]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-18). В 1983 году Майкл Дайер разработал в Йельском университете систему BORIS, которая имела сходство с работой Роджера Шанка и В. Г. Ленерта[[19]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0" \l "cite_note-19).

В третьем тысячелетии появились системы, использующие машинное обучение для классификации текста, такие как IBM [Watson](https://ru.wikipedia.org/wiki/IBM_Watson" \o "IBM Watson). Однако эксперты спорят о том, насколько «понимание» демонстрируют такие системы: *например*, по словам [Джона Серла](https://ru.wikipedia.org/wiki/%D0%A1%D1%91%D1%80%D0%BB,_%D0%94%D0%B6%D0%BE%D0%BD), Ватсон даже не понимал вопросов[[20]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-20).

Джон Болл, когнитивист и изобретатель Патомской теории[[21]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-21), поддерживает эту оценку. Обработка естественного языка позволила приложениям поддерживать производительность труда в сфере обслуживания и электронной коммерции, но это в значительной степени стало возможным благодаря сужению области применения. Существуют тысячи способов запросить что-то на человеческом языке, которые до сих пор не поддаются обычной обработке естественного языка. «Осмысленный разговор с машинами возможен только тогда, когда мы сопоставляем каждое слово с правильным значением, основанным на значениях других слов в предложении — точно так же, как трехлетний ребёнок делает это без догадок».

**Сфера и контекст**

Общий термин «понимание естественного языка» может применяться к разнообразному набору компьютерных приложений, начиная от небольших, относительно простых задач, таких как короткие команды, отдаваемые [роботам](https://ru.wikipedia.org/wiki/%D0%A0%D0%BE%D0%B1%D0%BE%D1%82), до очень сложных задач, таких как полное понимание газетных статей или отрывков из стихов. Многие реальные приложения находятся между двумя крайностями, например, [классификация текста](https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D1%8F_%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D0%BE%D0%B2) для автоматического анализа электронных писем и их маршрутизация в соответствующий отдел корпорации не требует глубокого понимания текста[[22]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-22), но требует с гораздо большим словарным запасом и более разнообразным синтаксисом, чем управление простыми запросами к таблицам базы данных с фиксированными схемами.

На протяжении многих лет предпринимались различные попытки обработки естественного языка или *англоподобных* предложений, представляемых компьютерам, с разной степенью сложности. Некоторые попытки не привели к созданию систем с глубоким пониманием, но повысили общее удобство использования системы. Например, Уэйн Рэтлифф изначально разработал программу *Vulcan* с синтаксисом, похожим на английский, чтобы имитировать англоязычный компьютер в «[Звездном пути](https://ru.wikipedia.org/wiki/%D0%97%D0%B2%D1%91%D0%B7%D0%B4%D0%BD%D1%8B%D0%B9_%D0%BF%D1%83%D1%82%D1%8C)». Позже Vulcan стал системой [dBase](https://ru.wikipedia.org/wiki/DBase" \o "DBase), чей простой в использовании синтаксис эффективно положил начало индустрии баз данных для персональных компьютеров[[23]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-23)[[24]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-24). Однако системы с простым в использовании синтаксисом, *похожим на английский*, сильно отличаются от систем, которые используют богатый [словарный запас](https://ru.wikipedia.org/wiki/%D0%9B%D0%B5%D0%BA%D1%81%D0%B8%D0%BA%D0%B0) и включают внутреннее представление (часто в виде [логики первого порядка](https://ru.wikipedia.org/wiki/%D0%9B%D0%BE%D0%B3%D0%B8%D0%BA%D0%B0_%D0%BF%D0%B5%D1%80%D0%B2%D0%BE%D0%B3%D0%BE_%D0%BF%D0%BE%D1%80%D1%8F%D0%B4%D0%BA%D0%B0)) семантики предложений естественного языка.

Следовательно, широта и глубина «понимания», к которому стремится система, определяют как сложность системы (и подразумеваемые проблемы), так и типы приложений, с которыми она может работать. «Широта» системы измеряется размерами её словарного запаса и грамматики. «Глубина» измеряется степенью, в которой его понимание приближается к пониманию носителя языка. В самых узких и неглубоких интерпретаторах команд, *похожих на английский язык*, требуется минимальная сложность, но они имеют небольшой диапазон приложений. Узкие, но глубокие системы исследуют и моделируют механизмы понимания[[25]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-25), но они все ещё имеют ограниченное применение. Системы, которые пытаются понять содержание документа, такого как выпуск новостей, помимо простого сопоставления ключевых слов, и оценить его пригодность для пользователя, шире и требуют значительной сложности[[26]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-26), но они все ещё не достаточно глубокие. Системы, которые являются одновременно очень широкими и очень глубокими, выходят за рамки современного уровня техники.

**Компоненты и архитектура**

Независимо от используемого подхода, большинство систем понимания естественного языка имеют некоторые общие компоненты. Системе нужен [словарь](https://ru.wikipedia.org/wiki/%D0%9B%D0%B5%D0%BA%D1%81%D0%B8%D0%BA%D0%B0) языка, [синтаксический анализатор](https://ru.wikipedia.org/wiki/%D0%A1%D0%B8%D0%BD%D1%82%D0%B0%D0%BA%D1%81%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7) и правила [грамматики](https://ru.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D0%BC%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0), чтобы разбивать предложения на внутреннее представление. Создание богатого словаря с подходящей [онтологией](https://ru.wikipedia.org/wiki/%D0%9E%D0%BD%D1%82%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%8F_(%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0)) требует значительных усилий, *например*, словарь [Wordnet](https://ru.wikipedia.org/wiki/WordNet" \o "WordNet) потребовал многих человеко-лет усилий[[27]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-27).

Системе также нужна теория из [*семантики*](https://ru.wikipedia.org/wiki/%D0%A1%D0%B5%D0%BC%D0%B0%D0%BD%D1%82%D0%B8%D0%BA%D0%B0_(%D0%B7%D0%BD%D0%B0%D1%87%D0%B5%D0%BD%D0%B8%D1%8F)), чтобы корректировать понимание. Возможности интерпретации системы понимания языка зависят от семантической теории, которую она использует. У конкурирующих семантических теорий языка есть определённые компромиссы в отношении их пригодности в качестве основы для автоматизированной компьютерной семантической интерпретации[[28]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-28). Они варьируются от *наивной семантики* или *стохастического семантического анализа* до использования [*прагматики*](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%B0%D0%B3%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0) для извлечения значения из контекста[[29]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-29)[[30]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-30)[[31]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-31). Семантические анализаторы преобразуют тексты на естественном языке в представления формального значения[[32]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-32).

Расширенные приложения для понимания естественного языка также пытаются включить логический [вывод](https://ru.wikipedia.org/wiki/%D0%92%D1%8B%D0%B2%D0%BE%D0%B4_(%D1%80%D0%B0%D1%81%D1%81%D1%83%D0%B6%D0%B4%D0%B5%D0%BD%D0%B8%D0%B5)) в свою структуру. Обычно это достигается путем преобразования производного значения в набор утверждений в [логике предикатов](https://ru.wikipedia.org/wiki/%D0%9B%D0%BE%D0%B3%D0%B8%D0%BA%D0%B0_%D0%BF%D0%B5%D1%80%D0%B2%D0%BE%D0%B3%D0%BE_%D0%BF%D0%BE%D1%80%D1%8F%D0%B4%D0%BA%D0%B0), а затем с помощью [логической дедукции](https://ru.wikipedia.org/wiki/%D0%94%D0%B5%D0%B4%D1%83%D0%BA%D1%82%D0%B8%D0%B2%D0%BD%D0%BE%D0%B5_%D1%83%D0%BC%D0%BE%D0%B7%D0%B0%D0%BA%D0%BB%D1%8E%D1%87%D0%B5%D0%BD%D0%B8%D0%B5) для получения выводов. Следовательно, системы, основанные на функциональных языках, таких как [Лисп](https://ru.wikipedia.org/wiki/%D0%9B%D0%B8%D1%81%D0%BF), должны включать подсистему для представления логических утверждений, в то время как системы, ориентированные на логику, такие как системы, использующие язык [Пролог](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%BE%D0%BB%D0%BE%D0%B3_(%D1%8F%D0%B7%D1%8B%D0%BA_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D1%8F)), обычно полагаются на расширение встроенной структуры логического представления[[33]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-33)[[34]](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-34).

Управление [контекстом](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BD%D1%82%D0%B5%D0%BA%D1%81%D1%82) при понимании естественного языка может представлять особые трудности. Большое разнообразие примеров и контрпримеров привело к множеству подходов к формальному моделированию контекста, каждый из которых имеет свои сильные и слабые стороны[[35](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BD%D0%B8%D0%BC%D0%B0%D0%BD%D0%B8%D0%B5_%D0%B5%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0#cite_note-35)