# Введение

В настоящее время бурное развитие компьютерной техники и информационных технологий привело к тому, что компьютер, мобильные устройства и информационные технологии присутствуют во всех сферах жизни современного человека. Компьютеры и мобильные устройства стали неотъемлемой частью современного мира. Развитие компьютерной техники с момента изобретения интегральной схемы резко ускорилось и происходит невероятно быстрыми темпами и на сегодняшний день. Столь же стремительно развивается и процесс миниатюризации компьютеров. Первые электронно-вычислительные машины были огромными устройствами, весившими тонны, занимавшими целые комнаты и требовавшими большого количества обслуживающего персонала для успешного функционирования. В контрасте с этим, современные компьютеры — гораздо более мощные и компактные и гораздо менее дорогие — стали воистину вездесущими.

Такое стремительное и быстрое развитие компьютеров и уменьшение их размеров приводит к тому, что привычные для нас способы взаимодействия человека с компьютером становятся все менее удобными. И перед производителями современных устройств и компьютерных систем возникла новая задача – создание нового способа человеко-компьютерного взаимодействия. Взаимодействие между пользователем и компьютером традиционно происходит с помощью различного программного и аппаратного обеспечения, например, с помощью образов или объектов на экране, или с помощью данных, полученных от пользователя посредством аппаратных устройств ввода (клавиатура, мышь).

Человеко-компьютерное взаимодействие – это целое научное направление, которое существует и развивается в целях совершенствования методов взаимодействия человека (пользователя) и компьютера. И основной задачей этого научного направления является улучшение этого самого взаимодействия, снижение барьера между человеческой моделью того, чего хотят достичь пользователи, и пониманием компьютера поставленных перед ним задач. И взаимодействие с компьютером с помощью экрана, мышки и клавиатуры на протяжении долгого времени отлично справлялось с этой задачей.

Но теперь вернемся к проблеме уменьшения размеров устройств. Уже сейчас на рынке пользуются большой популярностью «умные» часы, фитнесс-браслеты, миниатюрные музыкальные плееры, устройства «умного дома». Все эти устройства могут поместиться на ладони человека. И управление ими с помощью клавиатуры или мышки (пусть даже сенсорного экрана) трудно назвать удобным. По этой причине, в области человеко-компьютерного взаимодействия появился новый интерфейс – управление устройствами с помощью голосовых команд.

Идея управление компьютером с помощью голосовых команд за последнее время развилась в совершенно новую сферу информационных технологий, о которой раньше можно было лишь прочитать в книжках в жанре научной фантастики. И название этой новой области информационных технологий – обработка естественного языка (*Natural Language Processing, NLP*) или понимание естественного языка (*Natural Language Understanding, NLU*). То есть общение с компьютером и управление им с помощью естественной речи так, будто пользователь общается с таким же человеком.

# Актуальность темы

Как уже было сказано выше, понимание естественной человеческой речи – это совершенной новый способ взаимодействия человека и компьютера, который призван облегчить обмен информацией между пользователем и устройством и сделать его максимально удобным и привычным для человека.

Обработка естественного языка - совместное направление искусственного интеллекта и математической лингвистики. Оно изучает проблемы компьютерного анализа и синтеза естественных языков. Применительно к искусственному интеллекту анализ означает понимание языка, а синтез — генерацию грамотного текста. Решение этих проблем будет означать создание более удобной формы взаимодействия компьютера и человека. Подобными исследованиями занимается огромное количество компаний и небезуспешно. Достаточно вспомнить такие популярные голосовые помощники на мобильных устройствах, как Siri от корпорации Aplle, Cortana от Microsoft, Google Now от Google, Alexa от Amazon, и бесконечное количество чат-роботов, например в том же Telegram.

Алгоритм работы всех этих систем можно представить следующим образом:

1. Распознавание речи. В первую очередь необходимо распознать, что именно сказал пользователь, получить текстовую фразу из голосовой команды.
2. Анализ текста. На этом этапе из сказанной пользователем фразы необходимо извлечь информацию, проанализировать ее, проанализировать характер высказывания, тональность текста, извлечь «смысл» сказанного, понять, чего именно хочет пользователь.
3. Генерация текста. На основе анализа пользовательского текста необходимо сформировать какой-то ответ системы на запрос. Так как речь идет о взаимодействии с пользователем путем естественной речи, то и ответ компьютера тоже должен быть сформирован в виде текста на естественном языке.
4. Синтез речи. Компьютеру или мобильному устройству необходимо произнести подготовленный ответ.

Очевидно, что основной проблемой при разработке систем понимания естественной речи является понимание смысла сказанной человеком фразы и генерация грамотного, логичного ответа для пользователя. И решение этих проблем крайне актуально и необходимо, если мы и дальше хотим создавать новые, более компактные, удобные и умные устройства.

Поэтому, в своей работе я предлагаю решение одной из проблем, с которой сталкиваются разработчики систем понимания естественной речи, на этапе понимания сказанного пользователем текста и извлечения из него информации для дальнейшего анализа.

# Формулировка проблемы

В современных информационных технологиях роль такой процедуры, как извлечение информации, всё больше возрастает. А в такой научной области как понимание естественного языка процесс извлечения информации из сказанного пользователем текста играет ключевую роль. Примером извлечения информации может быть поиск некоторых информационных конструкций— формально это записывается так: *НанеслиВизит(Компания-Кто, Компания-Кому, ДатаВизита),* — из новостных лент, таких как: «Вчера, 1 апреля 2007 года, представители корпорации Пепелац Интернэшнл посетили офис компании Гравицап Продакшнз». Главная цель такого преобразования — возможность анализа изначально «хаотичной» информации с помощью методов обработки данных. Решением такой задачи занимается компьютерная лингвистика - научное направление в области моделирования интеллектуальных процессов при создании систем искусственного интеллекта, которое ставит своей целью использование математических моделей для описания естественных языков.

Процесс извлечения информации (смысла) из сказанного пользователем текста по сути сводится к извлечению из текста следующих смысловых сущностей:

1. Извлечение намерения пользователя (Intents). Намерения представляют собой отображение того, что сказал пользователь, и какие меры должны быть приняты вашим программным обеспечением.
2. Извлечение действия (Actions). Действия соответствуют шагам, которые ваше приложение будет предпринимать, когда пользователь выразит определенные намерения. Действие может иметь параметры для указания некоторой информации для приложения.
3. Сохранение и извлечения контекста (Contexts). Контекст представляет собой историю диалога с пользователем, которая позволяет точнее определить смысл текущей фразы и дифференцировать различные намерения и действия пользователя в зависимости от того, что было сказано ранее.

Извлечение из текста (произнесенной фразы) причисленных выше смысловых сущностей – задача всех систем, которые занимаются пониманием естественной человеческой речи. И этот процесс основывается практически на одних и тех же алгоритмах обработки данных. Рассмотрим принцип работы большинства таких систем и определим проблемы, с которыми они сталкиваются.

Процесс определения намерений пользователя и требуемых от вашего приложения действий основан на описании некоторых грамматических шаблонов и словарей для синтаксиса естественного языка. Рассмотрим этот алгоритм на примере следующей фразы:

*Can you describe witness in case 12?*

Для того, чтобы система могла определить смысл сказанной фразы, необходимо описать некоторые шаблоны, которые будут относиться к одному конкретному намерению пользователя, назовем это намерение (Intent) – *DescribeWitness*. Придумаем шаблон для конкретно этого намерения:

* @canyou @describe @witness @in @incident @number

Где символ «@» обозначает словарь, например, @canyou={can you tell, please tell me, give me, please, @null…}, @incident={case, incident, occasion…}, @describe={describe, give, show…}, @witness={witness, eyewitness, bystander} и так далее.

Таким образом, описав некий грамматический шаблон и заполнив определенные словари, из которых состоит этот шаблон, мы научили систему определять намерение *DescribeWitness* пользователя для довольно большого числа фраз. Под данный грамматический шаблон, очевидно, можно отнести следующие фразы: *please describe witness in incident 12, show me witness in case 12, show eyewitness in incident 12 и так далее.*

Получается, что системы, которые разрабатываются для понимания естественного языка, в своей работе используют описание грамматики языка в виде **словарей** и **шаблонов**, состоящих из этих словарей. И весь алгоритм понимания сказанной фразы заключается в том, чтобы определить – под какой именно грамматический шаблон попадает эта фраза, а уже после того, как будет определен шаблон, можно решить, какое намерение и действие несет в себе произнесенная фраза.