Лекция 7
«Проблемы понимания естественного языка»

Автор: д.т.н Джомартова Ш.А.

Проблемы понимания естественного языка

- Проблема СМЫСЛ-ТЕКСТ.
- *Проблема планирования* возникает при необходимости вести диалог
- Проблема равнозначности.
- Проблемы моделей участников общения. У участников общения должны быть сопоставимые *модели представления знаний*, необходимая глубина понимания, возможность *логического вывода*, возможность действия.
- Проблема эллиптических конструкций, то есть опущенных элементов диалога.
- Проблема временных противоречий

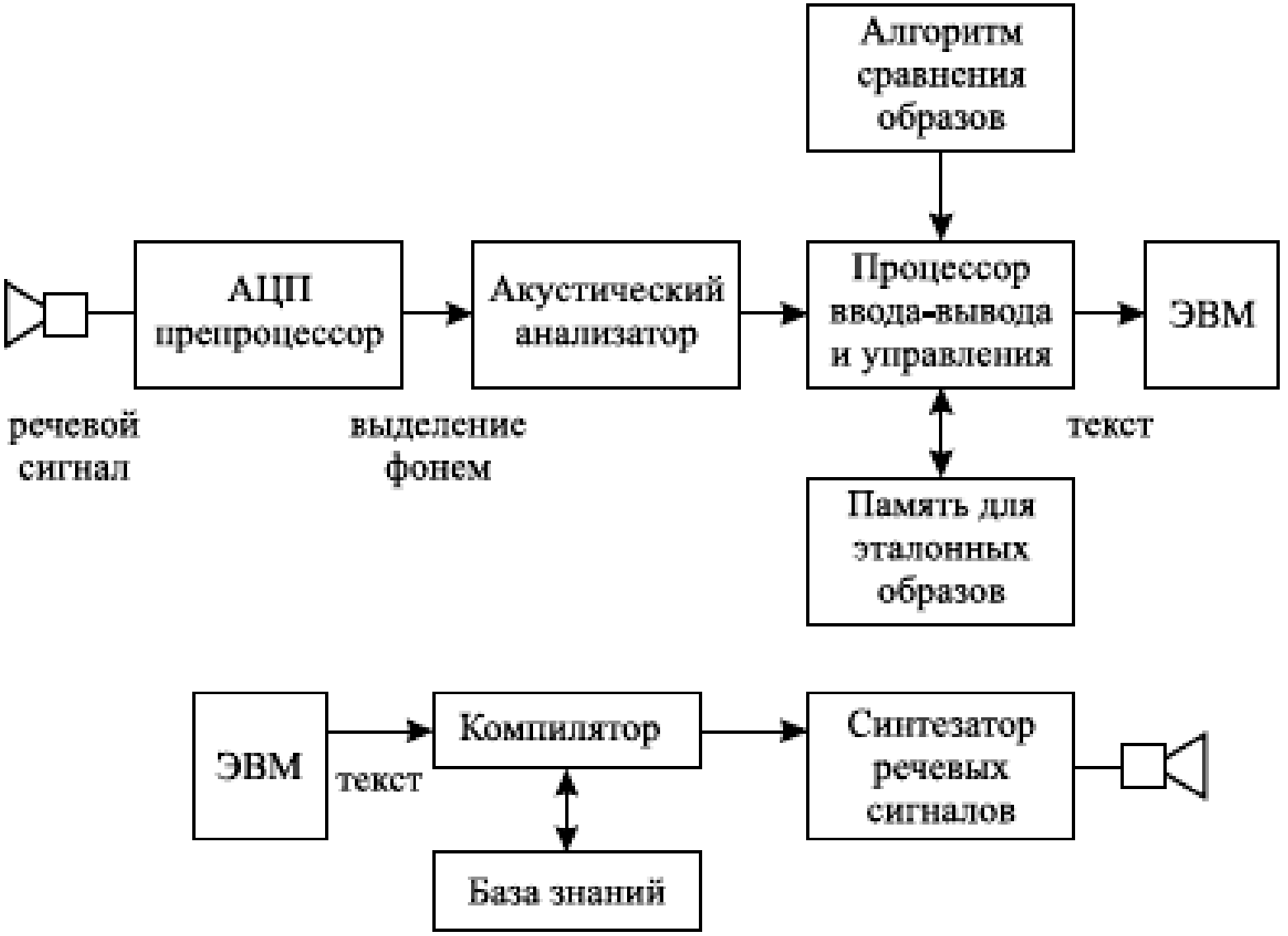
Основные понятия

- ***Слово*** - одна из основных единиц языка
- ***Предложение*** - любое высказывание
- ***Словосочетание*** - простейшая единица речи
- ***Дискурс*** - связный текст
- ***Лексема*** - слово во всей совокупности его лексических значений
- ***Морфема*** - минимальная законченная часть слова.
- ***Аффикс*** - прикрепленная к корню часть слова
- ***Омонимы*** - разные по значению, но одинаковые по написанию слова
- ***Синонимы*** - разные по написанию слова, но одинаковые по значению
- ***Эллипсис*** - опущенные слова в предложении
- ***Анафора*** - повторение объектов предложения

Основные понятия



Системы речевого общения



Обработка естественного языка (Natural Language Processing, NLP)

- Основные задачи АОТ:
  - Распознавание речи
  - Анализ текста
  - Извлечение информации
  - Информационный поиск
  - Анализ высказываний
  - Анализ тональности текста
  - Вопросно-ответные системы
  - Генерирование текста
  - Синтез речи
  - Задачи анализа и синтеза в комплексе
  - Машинный перевод
  - Автоматическое реферирование, аннотирование или упрощение текста


What is
Natural language processing
NLP


Машина отвечает аудиофайлом
Человек разговаривает с машиной
Как работает NPL
Преобразование данных в аудио
Машина захватывает аудио
Машина обрабатывает текст
Преобразование аудио в текст


ПРЕИМУЩЕСТВА
NLP
Позволяет персональным помошникам интерпретировать произнесенные слова
Позволяет использовать чат-боты для помощи клиентам
Автоматическое создание резюме большого и сложного текстового контента
Производит анализ настроения
Повышение эффективности и точности документирования
Расширенные аналитические данные, которые ранее были не доступны

ПРОБЛЕМЫ



Опечатки



Языковые раздичия



Врожденные предупреждения



Слова с несколькими значениями



Неопределенность и ложные срабатывания



Обучающие данные


Распознавание речи
Речевые теги
Смысл слова
Распознавание названного объекта
Разрешение совместной ссылки
Генерация естественного языка
ЗАДАЧИ NLP

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ



Обработка интеллектуальной литературы



Анализ настроения



Обнаружение мошеничества



Распознавание языка



Разговорный ИИ/Чат-бот



Обобщение текста



Перевод текста



Вопрос-Ответ



Редактирование данных



Социальный мониторинг СМИ



Бизнес аналитика



Медицинская документация

ОТРАСЛИ, ИСПОЛЬЗУЮЩИЕ NLP



Примеры NLP в жизни

- ***Email assistant:*** автокоррекция, проверка грамматики и орфографии, а также автозаполнение — все функции содержат NLP.
- ***Ответы на вопросы:*** если вы совершали покупки в Интернете или общались в чате на сайте, скорее всего, вы общались не с человеком, а с чат-ботом.
- ***Управляет e-commerce***: NLP обеспечивает лучшие результаты поиска при совершении покупок в Интернете.
- ***Извлечение и обобщение информации:*** NLP может извлекать и синтезировать информацию из различных текстовых источников

Примеры NLP в жизни

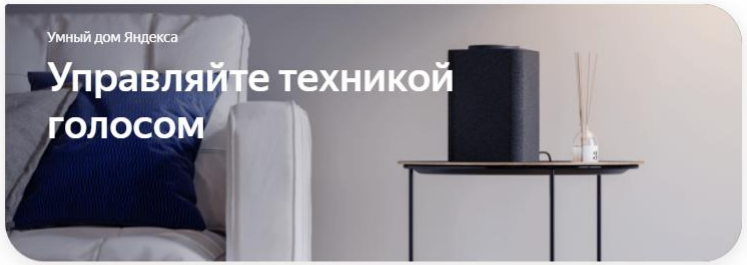
- NLP помогает приложению Livox быть средством связи для людей с ограниченными возможностями.
- Другим инструментом, содержащим технологию, является SignAll, который преобразует язык жестов в текст.
- Машинный перевод — это огромное приложение для NLP, которое позволяет нам преодолевать барьеры в общении с людьми со всего мира
- Технология NLP применяется для технического обслуживания самолетов.
- Технология NLP используется в предиктивной работе полиции

Алиса (голосовой помощник)

Подборки



Умный дом



Популярное

Чешские Тосты ★ 4
Навык для дружной компании. Слушайте тосты пана бармена из Велке-Поповице. Потребляйте ответственно — думайте...
Запусти навык Чешские Тосты

Примеры голосовых команд

Освещение
— Включи свет на кухне
— Выключи свет
— Убавь яркость света в коридоре

Розетка
— Включи розетку в ванной
— Выключи вентилятор
— Включи кофеварку на кухне

Как запустить

Скажите Алисе, что вы хотите приготовить и она расскажет вам, как это сделать. Например:

Как приготовить борщ | Давай приготовим шарлотку

Все | Завтрак | Обед | Ужин | Простое | Выпечка | Десерты | Быстрое | Посложнее | Праздни


Банановый хлеб
1 ч


Творожная запеканка
1 ч


Яйца Бенедикт
15 минут


Блины


Пирог с яблоками и корицей


Курица карри

Известия ★ 3,9
Я знаю много свежих новостей. Какая тема тебя интересует?
Запусти навык Известия

Новости Евразии
Экспертные оценки, классификация и анализ информации, касающейся ситуации в России и странах ближнего зарубежья...
Запусти навык Новости Евразии

День Победы ★ 4,2
Посвящается 75 летию победы. Хроника событий. 75 дней до победы. Навык можно запустить по фразе: День Победы
Запусти навык День Победы

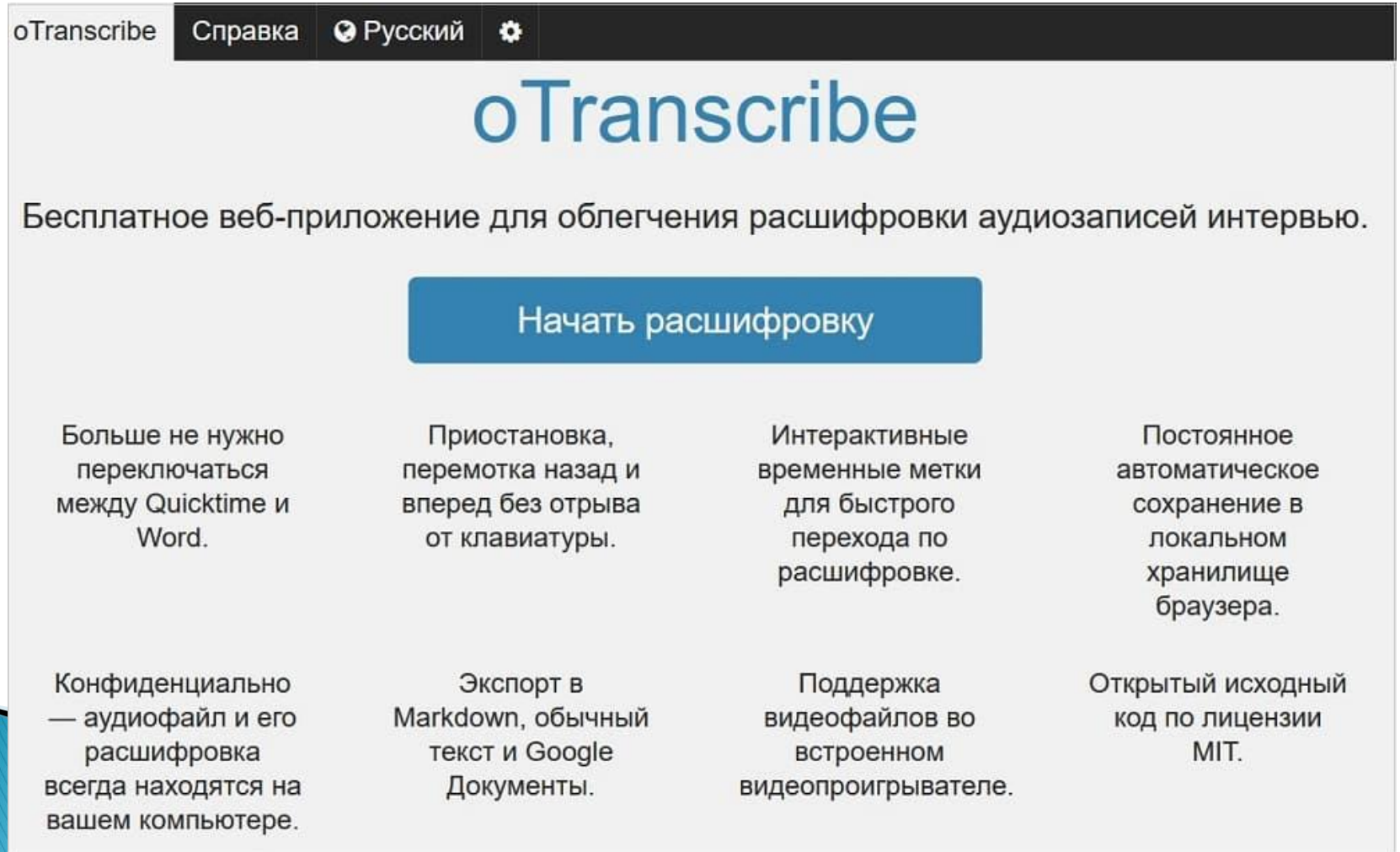
РИА Новости ★ 3,7
Новостная повестка от РИА Новости
Запусти навык РИА Новости

Сравнение голосовых помощников

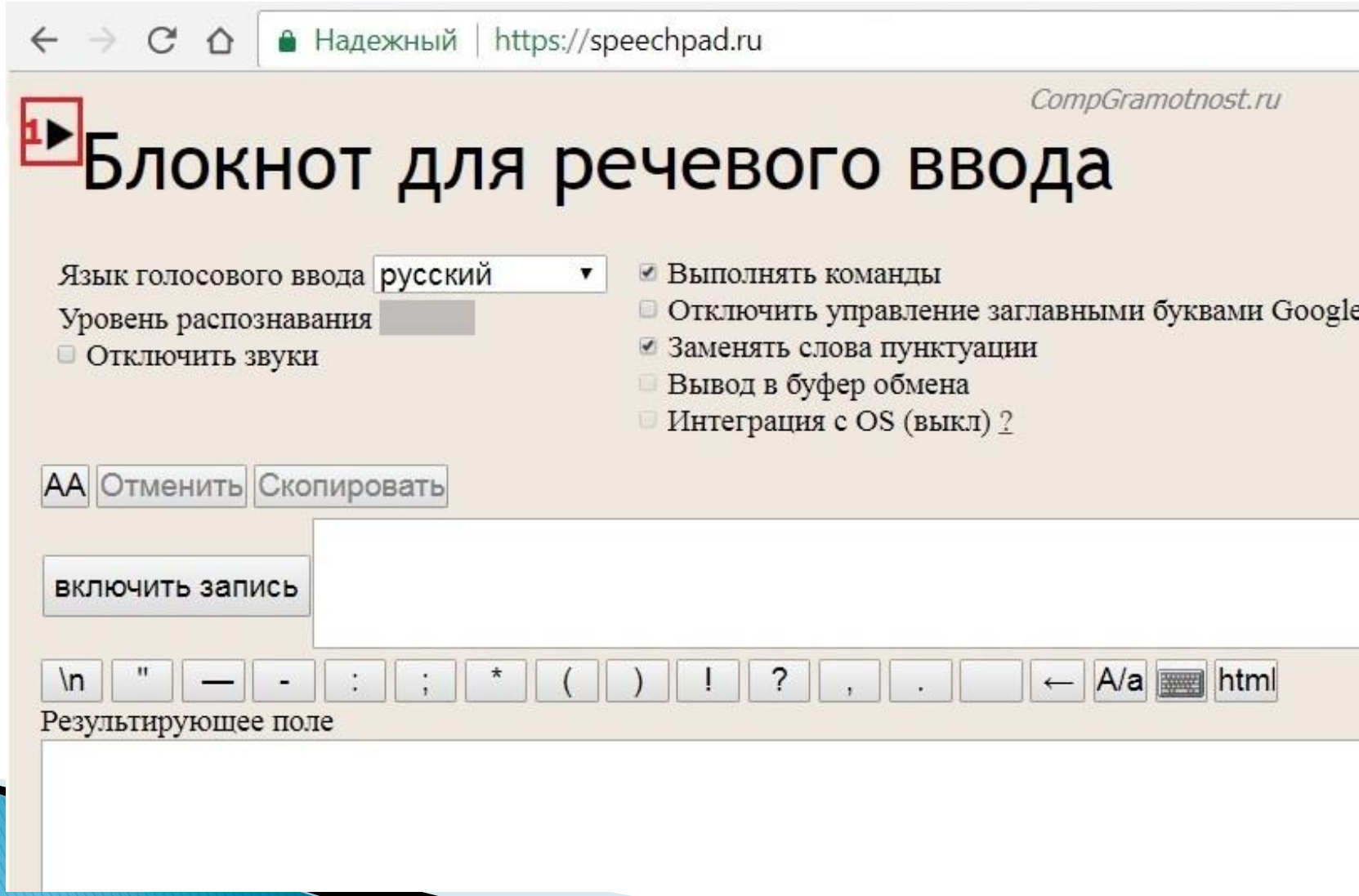
	Преимущества	Недостатки
Алиса	•говорит и шутит по-русски •работает с сервисами Яндекса •поддерживается бытовой техникой для российского рынка	приложение надо устанавливать отдельно не подойдет тем, кто больше пользуется сервисами Google или Apple
Ассистент Google	говорит по-русски поддерживает системы умного дома работает с сервисами Google установлен по умолчанию на гаджетах на Android	не подойдет тем, кто больше пользуется сервисами Яндекс или Apple
Siri	говорит по-русски поддерживает систему умного дома Apple работает с сервисами Apple установлен по умолчанию на гаджетах Apple	не подойдет тем, кто больше пользуется сервисами Яндекс или Google не работает на Android-устройствах
Маруся	говорит по-русски работает с сервисами Mail.ru	не подойдет тем, кто больше пользуется другим сервисами приложение надо устанавливать отдельно
Ассистент Дуся	понимает русский язык можно настроить функции на свой вкус поддерживает команды из Телеграм и с ПК	надо разбираться в настройках скриптов нужна платная версия (бесплатная работает только неделю)

Перевод речи в текст

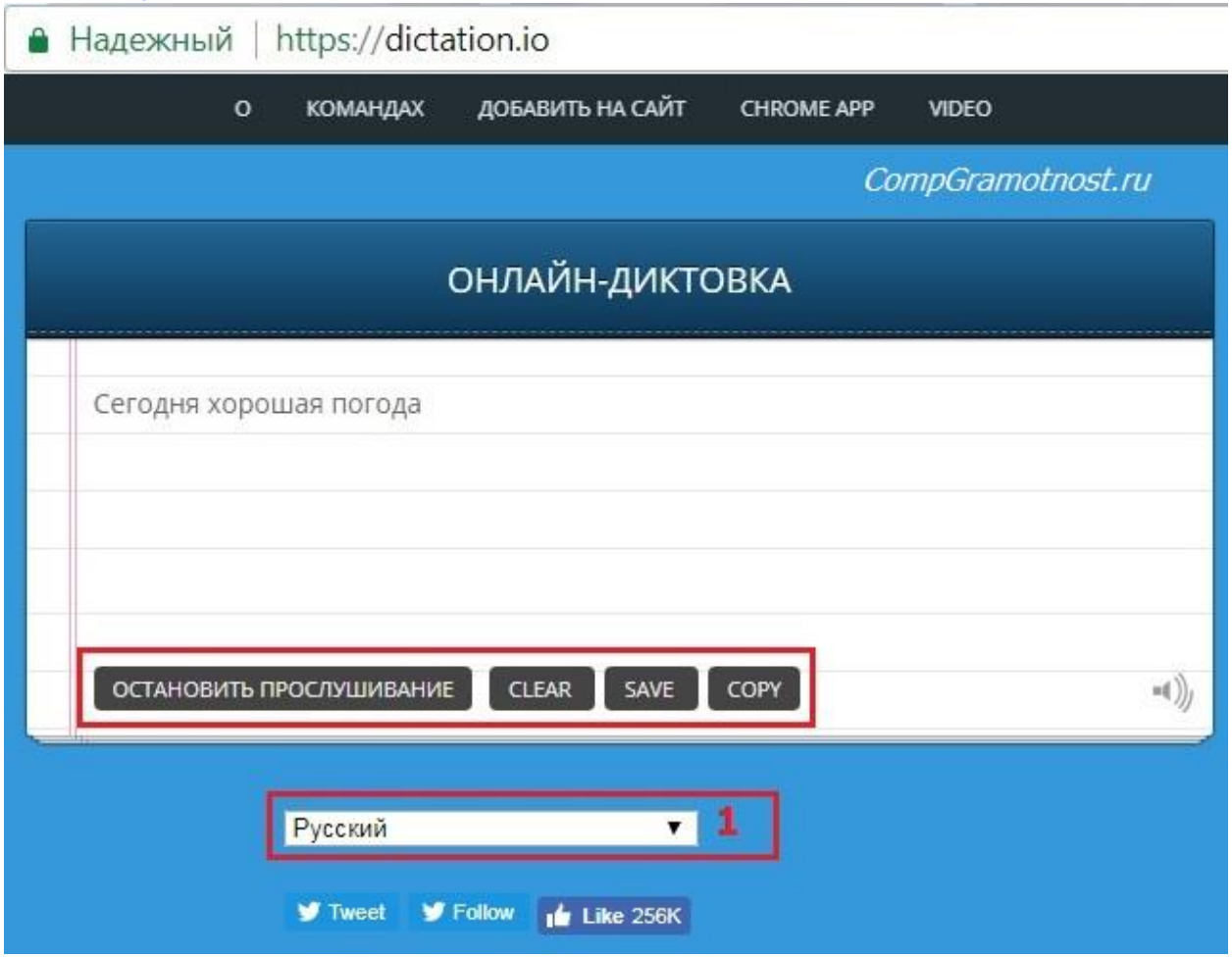
- *Транскрибация – это автоматический или ручной перевод речи в текст, точнее, запись аудио или видео-файла в текстовом виде.*



- Гугл Документы: Вы говорите, а компьютер записывает речь в виде текста
- Сайт speechpad.ru



- Сервис dictation.io



- RealSpeaker
- Speechnotes
- Dragon Dictation

10 лучших генераторов текстов на основе искусственного интеллекта (октябрь 2024 г.)





- https://www.jasper.ai/?fpr=uniteai
- Более 11,000 2,500 бесплатных шрифтов и категорий стилей письма.
- Поддерживает 30+ языков
- Интуитивно понятный интерфейс
- Помощник по написанию длинных форм (более 1,000 слов)
- Определите ключевые элементы в тексте (местоимения, глаголы, имена и т. д.)

Sudowrite

- https://www.sudowrite.com/?via=uniteai
- Sudowrite — это инструмент для письма на базе искусственного интеллекта, предназначенный для помощи писателям в их творческом процессе. Благодаря функциям мозгового штурма, инструментам написания, позволяющим создавать увлекательные романы за считанные дни, и инструментам редактирования, Sudowrite призван помочь писателям преодолеть творческие препятствия и разработать захватывающие сюжетные линии.

Scalenut

- https://www.scalenut.com/
- Scalenut особенно полезен для блоггеров и контент-маркетологов, которые регулярно публикуют объемный контент. Это комплексное решение со всем, что вам нужно для того, чтобы ваш контент занимал высокие позиции в поисковых системах, таких как Google, от эффективного исследования ключевых слов до полной SEO-оптимизированной статьи!



- https://surferseo.com/
- Surfer — это в первую очередь инструмент для создания SEO-контента, некоторые из основных функций которого включают в себя:
- Конструктор контуров .
- Тема открытия
- Объем ключевых слов и намерение поиска -
- Внутренняя структура контента
- Письмо ИИ
- ИИ-контент и плагиат



- https://writesonic.com/?via=uniteai
- Writesonic - это создание высококачественных копии веб-сайтов и маркетинговые тексты, технология создания контента на основе искусственного интеллекта
- Создает уникальный контент и широкий спектр тем за считанные секунды
- Встроенный инструмент исправления грамматики
- Генерирует контент для сообщений в блогах, объявлений, целевых страниц, описаний продуктов и многого другого.
- Перефразирует и расширяет содержание
- Поддержка 24 языки

BrandWell

- https://brandwell.ai/
- Разработанная для SEO и для веб-сайтов, которым необходимо масштабировать контент, генеративная модель ИИ предназначена для создания контента, похожего на человека, и проходит даже самые сильные и точные детекторы ИИ.

wordtune

- https://www.wordtune.com/
- Wordtune — один из самых передовых программных инструментов для написания ИИ на рынке. Эта технология позволяет создавать более убедительный и привлекательный контент с помощью передовой технологии НЛП на основе нейронной сети.
- После того, как вы предоставите Wordtune слова, он перефразирует содержимое и переписывает его, улучшая читаемость текста.

copy.ai

https://www.copy.ai/?via=uniteai

- Copy.ai предназначен для профессионалов SEO, при создании поста вы можете просто выбрать заголовок, ключевые слова, желаемый тон написания и цель статьи (например, обучение).
- Вы можете заставить его переписать абзацы и отшлифовать предложения. Затем просто скопируйте и вставьте работу в свою CMS для публикации.
- Выбирайте из 90 инструментов и шаблонов, чтобы быстро перейти к созданию контента. Варианты использования включают контент блога, списки и публикации в социальных сетях.

Copy Shark

- https://www.copyshark.ai/?via=antoine
- Copy Shark — это новый участник, который предлагает программное обеспечение на базе искусственного интеллекта, которое генерирует рекламные тексты, описания продуктов, коммерческие тексты, абзацы в блогах и многое другое. Он позволяет пользователям автоматически создавать уникальные и похожие на человека копии за считанные секунды, и, что самое главное, он поддерживает более 100 языков.

anyword

- https://www.anyword.com/
- Anyword — это инструмент для копирайтинга на основе данных, разработанный для маркетологов. Это позволяет создавать эффективные копии для рекламы, электронных писем, целевых страниц и контента для разных платформ.


JOIN THE FUTURE
OUR AI NEWSLETTER!
Subscribe Now
Ai



Спасибо за внимание!