

Использование технологий искусственного интеллекта для подготовки аудиокниг в формате DAISY

Михаил Пожидаев
Издательство «Директ-Медиа»

23 августа 2023 г.



Ключевая задача

Преобразование письменной речи в устную

Проблема номер один!

Для построения универсального алгоритма необходимы огромные обучающие данные, появление которых практически невероятно.

Решение!

Применение сочетания нейросетевых моделей с формальными правилами: формальные правила классифицируют языковые конструкции, а нейросети производят определение грамматических атрибутов.

Пример

Текст с добавленными аннотациями

На протяжении всего ТТСXIX в.ДЕВЯТНАДЦАТОГО
В+ЕКА наблюдалась трансформация понятия
ТТСдоговораДОГОВ+ОРА. Это объяснялось поднятием
цены с ТТС3 коп.ТРЁХ КОПЕЕК до ТТС1 руб.ОДНОГО
РУБЛЯ

Категории и технологии

Классификация языковых конструкций

1. Семантические омографы: модель Word2vec.
2. Грамматические (парадигмальные) омографы: модель BERT.
3. Склонение числительных: модель на основе долгой краткосрочной памяти.
4. Простейшие конструкции: формальные правила на основе недетерминированного конечного автомата.



www.directmedia.ru/

E-mail: mzp@luwrain.org

Группа в Телеграм: @DMGlagolitsa