

Семинар НУЛ учебных корпусов

19.01.2021

Портал для лингвистических исследований академических текстов

- Два типа текстов: эссе и проекты ВКР/статьи
- Три параметра оценки:
 - Грамматическая и пунктуационная корректность (Москва)
 - Сложность (Москва + Пермь в перспективе)
 - Стил (использование маркеров научного стиля)(Пермь)

Цели на этот год

- Добиться работоспособности нашего компонента портала
- Добавить в него модуль оценки
- Провести семинар по результатам исследований для АВС
- Начать работу над модулем для оценки сложности проектов ВКР и статей
- Расширить устный корпус
- Разработать список маркеров для устной речи
- Публикации, выступления на семинарах и конференциях, защиты курсовых и дипломных работ

Задачи на этот год (научные сотрудники)

- - составление списка маркеров научного стиля для устной речи (Смирнова Е.А., Туляков Д.С.);
- - разработка рекомендаций пользователям на основе полученной статистики использования маркеров в различных дисциплинах (Смирнова Е.А., Авраменко И.А, Туляков Д.С.);
- - валидация автоматического модуля поиска маркеров научного стиля на точность и полноту возврата (Смирнова Е.А., Авраменко И.А, Туляков Д.С.);
- - проектирование архитектуры Интернет-портала для лингвистических исследований, управляемого метаданными и допускающего динамическое расширение функциональных возможностей за счет подключения модулей расширения (Ланин В.В.);
- - расширение корпуса устной речи, его аннотирование (Смирнова Е.А., Авраменко И.А., Туляков Д.С.);
- - разработка компонента администрирования для системы лингвистических исследований (Ланин В.В.);
- - разработка системы оценочных метрик для описания проявления атрибутов стиля в тексте (Смирнова Е.А., Туляков Д.С., Авраменко И.А.);
- - внедрение модуля оценки качества текста в систему автоматизированного поиска маркеров научного стиля (Ланин В.В.)

Задачи на этот год (стажеры-исследователи)

- - проектирование и реализация системы оценки сложности текста на английском языке, администрирование сайта НУЛ (Иксанова А.А.);
- - разработка лексико-синтаксических шаблонов и программных модулей для идентификации маркеров научного стиля, работа над созданием устного корпуса (Гурова А.А.);
- - проектирование и реализация компонента статистических расчетов для систем лингвистических исследований, анализ статистической значимости различий между корпусами (Жуйкова С.К);
- - разработка компонента поиска в корпусе текстов для систем лингвистических исследований, работа над созданием устного корпуса (Салахов Р.Д.);
- - реализация унифицированного хранилища корпусов текстов, работа над созданием устного корпуса (Зайцев К.Н);
- - реализация компонентов системы для проведения независимого экзамена по английскому языку, работа над созданием устного корпуса (Охота Б.В.);
- - реализация компонентов системы для проверки эссе на английском языке, работа над созданием устного корпуса (Решетняк Р.М.);
- - изучение методов анализа лексической сложности текста, работа над созданием устного корпуса (Яркова В.В.).

Задачи на ближайшее время

- Темы на научных докладах на семинарах (а потом на конференциях) на первое полугодие 4 шт (к 26.01)
- Анализ статистики, разработка рекомендаций к семинару AWR – Смирнова Е.А., Туляков Д.С., Авраменко И.А. (к 26.01)
- Валидация системы поиска маркеров - Смирнова Е.А., Туляков Д.С., Авраменко И.А., В.В. Ланин (к концу февраля)
- Расшифровка и разметка докладов с конференции – Решетняк Р., Охота Б., Зайцев К., Салахов Р., Гурова А. (к концу мая)
- Работа над созданием общей платформы