

Программа лекций
« Корпусная лингвистика и цифровая гуманитаристика »
г-на Алексея Лаврентьева (Alexey Lavrentiev),

доктора филологических наук Нормальной школы г. Лион (École normale supérieure de Lyon), старшего научного сотрудника SNRS (Societe Nationale des recherches scientifiques) для студентов и преподавателей МГПУ

Время проведения: декабрь 2020 г.

Лекция 1. Гуманитарии и компьютерные технологии: непростая история взаимоотношений.

- Качественный и количественный анализ текстов как метод лингвистических и литературоведческих исследований до появления компьютеров
 - Грамматика Панини
 - Конкордансы в библеистике и теологии
 - Лоренцо Валла, «О подложности Константина дара» (1440)
 - Стилеметрия и «лингвистические спектры» Н.А. Морозова
 - Частотные словари
 - История языка
- Первые «компьютеризированные» гуманитарные проекты
 - Роберто Буза и Index Thomisticum
 - Брауновский корпус
 - Математическая лингвистика в СССР
 - Корпус «СКАТ»
 - Проект Машинного фонда русского языка
- На пороге цифровой революции (1980-е – 1990-е годы)
 - Компьютерные и диванные лингвисты
 - Национальные корпусы
 - Рост числа и разнообразия ресурсов
- Движение «цифровой гуманитаристики»
 - Зарождение и история
 - Организации и конференции
 - Основные направления исследований
 - Анализ текстов, социальных сетей
 - Мультимедийные ресурсы
 - ГИС
 - Семантическая сеть
 - Междисциплинарные исследования

Лекция 2. Текстометрия и «дистантное чтение» на базе платформы TXM

- Концепция «дистантного чтения»
- Текстометрия: методы и технологии
- Качественный анализ
 - Конкорданс
 - Гипертекстовая навигация
 - Прогрессия
 - Язык запросов CQL
- Количественный анализ
 - Частотные словари
 - Подкорпусы и разбивки
 - Специфичность

- Факторный анализ
- Текстометрический анализ корпуса русского рассказа первой трети XX в.

Лекция 3. Создание и анализ текстовых и мультимедийных корпусов. Корпусные технологии в лингводидактике

- Использование корпусов в лингводидактике
 - Корпус как дополнение к словарю
 - Корпуса ошибок
- Создание корпуса
 - Форматы текстовых данных и аннотации, TEI XML
 - Виды разметки
 - Автоматическая обработка текста
 - Импортирование и анализ корпуса в TCM