**Проектное предложение**

|  |  |
| --- | --- |
| Тип проекта | *Прикладной* |
| Название проекта | корпус классической персидской поэзии |
| Подразделение инициатор проекта | ИКВИА |
| Руководитель проекта | Арманд Елена Евгеньевна |
| Заказчик проекта / востребованность проекта | поэтическим корпусом смогут пользоваться не только исследователи персидской поэзии, персидского языка и литературы, но и специалисты по поэтике, а также лингвисты, использующие в своих исследованиях языковые корпуса |
| Основная проектная идея / описание решаемой проблемы | Проект предполагает ручную и полуавтоматическую работу с разметкой персидского поэтического корпуса. Нужно будет проверять корректность морфологической разметки, находить системные ошибки в присвоении грамматических тегов и предлагать решения по исправлению этих ошибок. |
| Цель проекта | Нужно будет либо работать с поисковым интерфейсом корпуса, либо с выгруженными разработчиками из корпуса списками слов с их леммами и грамматическими тегами. В корпусе имеется класс низкочастотных слов, для которых не получилось сделать удовлетворительной морфологической разметки, нужно будет либо разметить их вручную, либо предложить формальные описания для таких слов, которые позволили бы доразметить эти слова автоматически. |
| Планируемые результаты проекта, специальные или функциональные требования к результату | 1) Пополнение корпуса новыми текстами.  2) Ревизия грамматической разметки.  3) Снабдить тексты метрической разметкой и информацией о рифме и редифе. |
| Требования к участникам с указанием ролей в проектной команде при групповых проектах | Знание современного или классического персидского языка. |
| Количество вакантных мест на проекте | 25 |
| Проектное задание | Провести ревизию системы грамматической разметки корпуса. Возможно, потребуется перераспределить грамматические формы внутри грамматических категорий и сделать всю систему разметки более удобной и конвенциональной.  Перспективной работой была бы подготовка машиночитаемого двуязычного словаря, которая позволила бы показывать в выдаче текстов в корпусе не только грамматическую разметку и леммы, но и перевод отдельных слов.  Наконец, большая часть текстов в корпусе пока остаётся без метрической характеристики. Полезной работой была бы ручная разметка текстов в терминах аруза и помощь в подготовке машинного инструмента определения метра стиха. |
| Критерии отбора студентов | знание современного или классического персидского языка |
| Сроки и график реализации проекта | сентябрь 2020 – июнь 2021 |
| Трудоемкость (часы в неделю) на одного участника | 2 часа |
| Количество кредитов | 3 |
| Форма итогового контроля | экзамен |
| Формат представления результатов, который подлежит оцениванию | Студенты будут поделены на группы, перед каждой группой будут поставлены определенные задачи. Будет оцениваться объем и качество проделанной работы. |
| Образовательные результаты проекта | Студенты научатся работать с корпусом, овладеют практическими знаниями персидского аруза (правилами стихосложения), смогут применить на практике знание грамматики классического персидского языка. |
| Критерии оценивания результатов проекта с указанием всех требований и параметров | 0.25 *–* определение метра  0.25 – словарная работа  0.5 – морфологическая разметка |
| Возможность пересдач при получении неудовлетворительной оценки | *Да* |
| Рекомендуемые образовательные программы | Язык и литература Ирана, Фундаментальная и компьютерная лингвистика |
| Территория | Москва, ФГН |