

**О коррекционных мерах, принятых в связи с недочётами, выявленными в
первичной версии технического задания на выполнение научно-исследовательского
проекта по теме:**

**«Семантический анализ текстов научных публикаций (разработка методов
семантического анализа текстов научных публикаций с целью создания на их основе
базы результатов научных исследований)»**

(пояснительная записка на 2 листах)

1. Сведения о действиях предпринятых в связи с комментариями членов НЭС

В отношении замечания №1

Текст замечания

В тексте нет каких-либо указаний на то, что имеется в настоящий момент в обозначенной области. Это крайне затрудняет понимание ценности проекта.

Ответ группы исполнителей

Структура технического задания не подразумевает указанного раздела, при этом по заявленной тематике имеется значительный объем результатов полученных в ходе работ факультета по смежной тематике поддержанных грантом Министерства образования и науки Российской Федерации в рамках Соглашения о предоставлении субсидии от «26» сентября 2017 г. № 14.576.21.0092 (Уникальный идентификатор соглашения RFMEFI57617X0092) на выполнение прикладных научных исследований по теме: «Разработка нейросетевой системы прогнозирования авиапроисшествий и управления рисками безопасности полетов на основе ретроспективных данных, включающих множество параметров и текстовых описаний событий».

В отношении замечания №2

Текст замечания

Из текста ТЗ не понятно, создается ли это только для русскоязычных статей или может использоваться и для других языков.

Ответ группы исполнителей

Подход может использоваться для различных языков, в ходе заседания НЭС исходные работы будут проводиться для английских текстов.

В отношении замечания №3

Текст замечания

Если отчет промежуточный, то не ясно, когда будет представлен его результат - не отражено в плане.

Ответ группы исполнителей

В план добавлен пункт о представлении промежуточных результатов в октябре текущего года.

В отношении замечания №4

Текст замечания

**Семантический анализ текстов научных публикаций (разработка методов семантического анализа текстов научных публикаций с целью создания на их основе базы результатов научных исследований)
(пояснительная записка)**

1. Сократить название, возможно, так: "Создание базы результатов научных исследований на основе семантического анализа научных текстов"
2. Исправить многочисленные пунктуационные ошибки (см. пункты ТЗ №№ 1, 8, 9, 14), а в пункте 8 слово "также" написать слитно..

Ответ группы исполнителей

Для разрешения существующих терминологических противоречий предлагается исправить название темы на более однозначное

Разработка программных средств автоматической обработки научных текстов с использованием моделей компьютерной лингвистики для формирования базы результатов научных исследований

Найденные при редактировании ТЗ ошибки исправлены

В отношении замечания №5

Текст замечания

Вопросы: 1) В новизне: "Получение комплекса инструментальных средств для автоматизированной оценки результатов научных исследований, представленных в форме научных публикаций...". Итак, результаты научных исследований все-же будут оцениваться через семантику текстов? Как? Вы же не согласны с подобным замечанием. А "отношение автора к предмету исследования" может сказать что-то о качестве его научных результатов?

2) Что понимается под "латентной структурой связей и отношений между исследования представленными в форме научных статей" и что она даст конкретному студенту, магистранту, аспиранту, который зашел в вашу базу знаний и просто хочет найти научные статьи по своей теме исследования? Хотелось бы, чтобы эти статьи были высокого качества, но по каким критериям (вы их называете "маркерами дискурса"?). Если критерии только формальные, я еще могу с этим согласиться. Но если критерии содержательные, то вы же не специалисты в психологии. Кроме того, эти критерии одни и те же для статей по медиевистике в декоративно-прикладном искусстве народов Востока и по современным методам добычи и переработки в горно-рудной промышленности? Не сочтите это за язвительность, но пока у меня вопросов в вашем проекте значительно больше, чем ответов.

3) Кстати, если это еще одна поисковая система, то зачем она нужна, если есть базы WOS, Scopus, elibrary и многие другие? Будет ли она конкурентоспособна с ними? Просто потому что она будет принадлежать университету (что само по себе совсем неплохо)?

Ответ группы исполнителей

**Семантический анализ текстов научных публикаций (разработка методов семантического анализа текстов научных публикаций с целью создания на их основе базы результатов научных исследований)
(пояснительная записка)**

Для снятия противоречий связанных с толкованием термина «семантический анализ», переработано название проекта, изменён пункт 2, в котором указывается практическая проблема на решение которой направлена работа, в новой редакции он звучит следующим образом

Снижение времени необходимого для анализа актуального состояния научных исследований в определённой предметной области. Результат достигается за счёт представления доступных публикаций по референтной теме в форме интерактивного графа в трёхмерном пространстве, спроецированном на плоскость. При этом узлы графа представляют выделенную в ходе автоматического анализа характеристику научного текста (результат исследования, методы исследования и др.), а рёбра отождествлены с отношениями между этими характеристиками.

Дать более детальный ответ на замечание сложно т.к. оно подразумевает что планируется проведение работ принципиально отличающихся от заявленных в ТЗ

В отношении замечания №6

Текст замечания

Какие именно тексты будут анализироваться и по каким специальностям, к сожалению, мало что поняла по-прежнему из ТЗ.

Ответ группы исполнителей

Ввиду специфики методов, используемых в компьютерной лингвистике, вопрос о специальностях не является принципиальным, т.к. предлагаемые в ТЗ программные разработки носят универсальный характер, конкретный перечень работ для промежуточной демонстрации результатов исследования будет определён на первом этапе работ, в случае если проект получит поддержку.

В отношении замечания №7

Текст замечания

1. не совсем понятна основная цель проекта. Вы пишете про формирование подхода, в то же время в новизне речь идет о получении "инструментальных средств" (может быть, все-таки просто "инструмента"?) для автоматизированной оценки результатов исследования. Формирование подхода и разработка инструмента - две качественно разные задачи.

2. слишком длинное название предмета исследования. Советую переформулировать.

3. желательно писать текст, используя четкие понятия и формулировки. Наличие большого количества абстрактных предложений, за которыми не стоит глубокий смысл, раздражает и отвлекает внимание от главного.

**Семантический анализ текстов научных публикаций (разработка методов семантического анализа текстов научных публикаций с целью создания на их основе базы результатов научных исследований)
(пояснительная записка)**

Ответ группы исполнителей

В формулировки внесены изменения детализирующие предлагаемые работы, название предмета исследования сокращено.

В отношении замечания №8

Текст замечания

Не понятен адресат столь сложного в техническом отношении исследования при обилии поисковых систем

Каким образом возможно "автоматическое извлечение смысловых единиц"?

Ответ группы исполнителей

Предлагаемые работы не дублируют функции поисковых систем, хотя ряд методов распространённых в компьютерной обработке запросов на естественном языке будет использоваться при реализации проекта

Ответ на второй пункт замечания не укладывается в формат пояснительной записки и требует строгой договорённости снимающей имеющуюся лексико-семантическую контекстную неопределенность для ряда понятий использованных при написании ТЗ