ТЕХНИКАЛЫҚ ҒЫЛЫМДАР
ТЕХНИЧЕСКИЕ НАУКИ
TECHNICAL SCIENCE

CENTRAL ASIAN SCIENTIFIC JOURNAL
ISSN 3135-3061
2026. №4 (32), Том 2. С. 127–134
DOI: https://doi.org/10.5281/zenodo.23837227
Сайт журнала: https://cajournal.kz

УДК 004.93:378.146

ПОДБОР ТЕМ ДЛЯ ИССЛЕДОВАНИЙ И РАСПРЕДЕЛЕНИЕ РУКОВОДИТЕЛЕЙ С УЧЕТОМ ОГРАНИЧЕНИЙ ПО РЕСУРСАМ

С.Ж. Советов[1], Д.С. Байгожанова[2]
[1]Международный Университет Астана, Астана, Казахстан
[2]Международный университет Астана, Астана, Казахстан

Советов Султан Жамбулович — Студент 2го магистратуры, Кафедра информационных технологий и инженерии, Международный Университет Астана, Астана, Казахстан.

Байгожанова Даметкен Сагидуллаевна — к.п.н., ассоциированный профессор, Высшая школа информационных технологий и инженерии, Международный университет Астана, Астана, Казахстан.

© С.Ж. Советов, Д.С. Байгожанова, 2026

Аннотация: Подбор научных проектов для научных руководителей требует как поиска с учетом языка, так и четкой политики распределения. В данном исследовании эти этапы реализуются и оцениваются отдельно с использованием диагностического корпуса: 24 варианта предложений на английском, русском и казахском языках и 24 англоязычных профиля научных руководителей. Метки релевантности, полученные на основе таксономии, отличаются от независимых экспертных оценок. TF-IDF и BM25 сравниваются с многоязычным MiniLM в формате unsigned-int8 ONNX и многоязычным E5-small в формате float32. Четыре алгоритма распределения оцениваются на 30 синтетических группах и в трех режимах пропускной способности. E5 достигает показателя макро nDCG@5 0,792 по сравнению с 0,685 для MiniLM, в то время как MiniLM имеет лучшие описательные рейтинги по запросам на английском и русском языках. Все 1440 распределений соответствуют ограничениям пропускной способности. Отложенное принятие не приводит к образованию блокирующих пар, но жертвует сконструированной релевантностью по сравнению с назначением максимального веса при независимо отобранных предпочтениях научных руководителей. Равномерная нагрузка при ограниченных возможностях возникает механически и не обеспечивает справедливости. Вклад заключается в поддающейся проверке экспериментальной реализации и диагностике с учетом языковых и ресурсных особенностей, а не в новом алгоритме сопоставления или доказательствах институциональной эффективности.

Ключевые слова: многоязычный поиск; темы исследований; распределение руководителей; отложенное принятие; ограничения по ресурсам; воспроизводимость.

Университетская платформа для выбора тем исследований и научных руководителей должна сопоставлять предложенное студентом исследование с соответствующей экспертизой и учитывать ограниченные возможности научного руководства. Это разные решения: подходящая рекомендация может стать недоступной после распределения, в то время как стабильное назначение может отражать предпочтения, которые не обеспечивают максимального соответствия исследованию. Данное исследование позволяет проверить это различие в контексте академического подбора научных руководителей.

В данном исследовании оценивается эффективность сопоставления многоязычных данных в контролируемых условиях с использованием лексического и векторного поиска,

распределения ресурсов с ограничениями по объему, явных меток ссылок и архивных промежуточных результатов.

В эксперименте сравниваются методы поиска информации по запросам на английском, русском и казахском языках, а также исследуется, как политика распределения ресурсов и пропускная способность влияют на релевантность, охват, стабильность и распределение рабочей нагрузки.

Связанные работы

Реймерс и Гуревич (2019) представили Sentence-BERT для получения многократно используемых представлений предложений, а их многоязычный подход к дистилляции выравнивает переводы в общем пространстве (Реймерс и Гуревич, 2020). Ванг и др. (2024) описывают многоязычное семейство E5, интерфейс поиска которого различает запросы и фрагменты текста. Эти методы мотивируют семантическое сопоставление описаний, различающихся по словарю или языку. Они не гарантируют, что каждый поддерживаемый язык будет одинаково эффективен при обработке исследовательских предложений.

Оценка MTEB показывает, что эффективность внедрения зависит от выполняемой задачи, а не от единого универсального ранжирования моделей (Muennighoff et al., 2023). В недавних работах конкретно изучалось сопоставление навыков и концепций на казахском, русском и английском языках с использованием адаптации предметной области (Serikbayeva et al., 2026). Это приложение связано с сопоставлением экспертных знаний, но его результаты не могут быть напрямую перенесены на выбор руководителя. Поэтому в настоящем сравнении представлены диагностические данные, специфичные для каждого языка, без тонкой настройки или предположения, что опубликованные эталонные показатели предсказывают результаты работы учреждения.

Исследования в области распределения ресурсов уже объединяют двусторонние предпочтения, квоты на надзор и соображения, касающиеся рабочей нагрузки. Санчес-Ангикс и др. (2019) исследуют многоцелевой эволюционный подход, в то время как Рамотсиси и др. (2022) формулируют модель целочисленной оптимизации в тематическом исследовании инженерного факультета. Следовательно, объединение сопоставления текста с ограничениями по пропускной способности не претендует на звание алгоритмического новшества. Вклад здесь заключается в воспроизводимом разделении качества представления, заявленной стабильности предпочтений и полезности назначения.

Методология

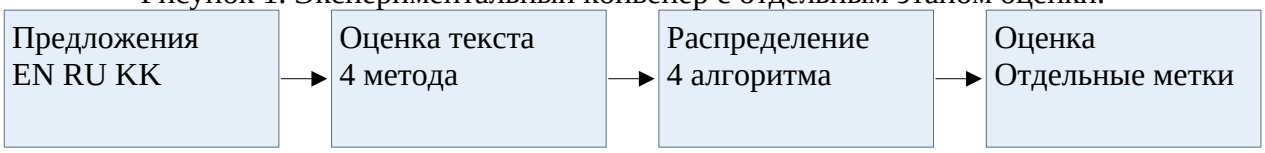
Созданный корпус и метки ссылок

Корпус содержит 12 тем исследований в области вычислительной техники, сгруппированных в шесть смежных пар: поиск и перевод; обнаружение вторжений и обучение с сохранением конфиденциальности; аналитика обучения и образовательные рекомендации; обучение представлению графов и графы знаний; временные ряды датчиков и периферийное зрение; стабильное сопоставление и составление расписаний. Два разных варианта формулировки для каждой темы дают 24 варианта. Каждый вариант имеет согласованную формулировку на английском, русском и казахском языках, что позволяет получить 72 запроса. Каждая тема также содержит два описания руководителей, что позволяет получить 24 профиля кандидатов на английском языке.

Для каждой из 1728 пар запрос-кандидат релевантность равна 2, если основные темы совпадают, 1, если они принадлежат к одной и той же смежной паре, и 0 в противном случае. Это сконструированные эталонные метки, а не экспертные эталонные метки. Каждый запрос имеет два кандидата 2-го уровня и два кандидата 1-го уровня. Рубрика смежности намеренно рассматривает широкую взаимосвязь как более слабую релевантность; это предположение, которое будущие эксперты могут отклонить или уточнить. Метки и идентификаторы тем используются только оценщиком.

© Central Asian Scientific Journal, 2026

Рисунок 1. Экспериментальный конвейер с отдельным этапом оценки.



Перевод намерения осуществляется путем повторных измерений, а статистический анализ объединяет связанные наблюдения на уровне темы.

Оценка и распределение с ограничениями

TF-IDF использует униграммы и биграммы слов, подобранные по документу, сублинейную частоту терминов и косинусное сходство. BM25 использует токены слов Unicode, состоящие как минимум из двух символов, k_1 = 1,5 и b = 0,75, с положительной логарифмически сглаженной обратной частотой документа (Robertson & Zaragoza, 2009). Термины запроса суммируются в отсортированном порядке для обеспечения воспроизводимости. Ни одна из лексических баз не переводит запросы, не стеммирует слова и не удаляет стоп-слова. Поэтому они проверяют прямой лексический перенос, а не максимально возможный многоязычный лексический конвейер.

Для получения плотных векторов используются методы paraphrase-multilingual-MiniLM-L12-v2, использующий официальный экспорт unsigned-int8 AVX2 ONNX, и multilingual-e5-small, использующий официальный экспорт float32 ONNX. Оба метода дают 384-мерные векторы. После пулинга с маской внимания выполняется L2-нормализация. E5 получает необходимые префиксы «query:» и «passage:». Все незамаскированные специальные токены участвуют в пулинге. Матрица сходства представляет собой скалярное произведение нормализованных векторов запроса и профиля:

Предпочтения студентов сортируют каждую строку по убыванию степени сходства, при этом идентификатор научного руководителя разрешает споры при совпадении результатов. Научные руководители получают независимые выборки строгих рейтингов студентов. Эти рейтинги представляют собой контролируемый источник двусторонних разногласий; они не являются реальными предпочтениями, оценками академической успеваемости или предполагаемой релевантностью. Допустимы все пары, и каждый участник предпочитает выполнимое задание, чем остаться без пары. Не применяется непроверенный порог семантического отклонения.

Четыре стратегии получают одинаковые входные данные. Случайное распределение перемешивает доступные места и студентов. Глобальный жадный алгоритм принимает ребра с уменьшающимся баллом, когда студент свободен и остается свободная вместимость. Отложенное принятие по методу Гейла-Шэпли, предлагающее студента, позволяет каждому руководителю сохранить наиболее предпочтительных кандидатов до достижения вместимости при наличии гибких предпочтений (Гейл и Шапли, 1962). Точное распределение с максимальным весом расширяет вместимость на одинаковые места и решает задачу линейного суммирования. Оно максимизирует сначала мощность множества, а затем сумму баллов сходства:

Здесь n - размер когорты, а K определяет максимально допустимую мощность множества. Точный решатель не использует метки релевантности и не накладывает ограничений на стабильность. Его оптимальный показатель зависит от входной модели; суммарные значения сходства между моделями не являются напрямую сопоставимыми. Блокирующая пара состоит из студента, предпочитающего другого научного руководителя, у которого есть вакантное место, или предпочитающего этого студента текущему

назначенному специалисту. Отложенное принятие гарантирует стабильность при заданных предпочтениях, а не максимальную семантическую полезность или сбалансированную нагрузку.

Реализация и экспериментальная установка

Реализация выполняется на Python 3.12.14 под управлением Windows 11 с процессором AMD64. В записанной среде используются NumPy 2.5.4, SciPy 1.18.1, scikit-learn 1.9.1, ONNX Runtime 1.31.1 и токенизаторы 0.23.3. Вывод на ЦП осуществляется с использованием одного потока внутри операции и одного потока между операциями, размером пакета 8 и фиксированной длиной с заполнением 128. Текст не обрезается: максимальная наблюдаемая длина составляет 33 токена MiniLM и 35 токенов E5. Точные ревизии репозитория и хеши файлов прилагаются к коду.

Каждая модель проходит один полный подготовительный этап и три прохода по всем 96 текстам с ограничением по времени. Второй полный эксперимент выполняется в другом процессе с другим хеш-параметром Python. Загрузки, загрузка модели и кодирование записываются отдельно. Перед оценкой проводятся тесты метрик, рассчитанные вручную, 50 небольших случайных случаев стабильности/оптимальности, исчерпывающее перечисление назначений и тесты с нулевой пропускной способностью. Каждое зарегистрированное распределение проверяется на нарушения квот, максимально возможное покрытие, стабильность отложенного принятия и верхнюю границу максимального балла.

В отчете о ранжировании nDCG@5 используется коэффициент усиления $2^r - 1$, где r - градуированная релевантность, и идеальный порядок для всех 24 кандидатов (Järvelin & Kekäläinen, 2002). MRR использует первого кандидата с $r > 0$ из всего списка; Precision@5 и Recall@5 также рассматривают $r > 0$ как положительный результат. Поскольку положительных результатов всего четыре, Precision@5 не может превышать 0,8. Интервал для каждого среднего значения использует 5000-процентные бутстрап-выборки 12 тематических кластеров, сохраняя как намерения, так и их переводы.

Для всех запросов и каждого языка оцениваются три метода сравнения с использованием всех 2^{12} шаблонов знаков на уровне тем, с поправкой Холма для полученных 12 тестов. В соответствии с проблемой выбора корректных тестов, обсуждаемой Дрором и др. (2018), эти тесты интерпретируются с учетом взаимозаменяемости на уровне тем.

Распределение осуществляется с использованием 30 фиксированных начальных значений. Каждая когорта отбирает 48 экземпляров шаблона с заменой, при этом вероятности тем пропорциональны и язык являются единообразными. Жесткие квоты обеспечивают 48 мест, избыточные квоты - 60, а дефицитные - 36. В общей сложности получается 4 оценщика × 4 политики × 3 режима × 30 когорт = 1440 распределений. Нормализованная по когортам релевантность составляет $\Sigma r/(2 \times 48)$, при этом для несовпадающих студентов она равна нулю. Архивируются данные об охвате, блокирующих парах, ранге назначенного студента, сожалении по оценке модели и коэффициенте Джини как нагрузки, так и использования мощностей. Парные интервалы распределения позволяют повторно выбрать 30 сценариев 5000 раз.

Результаты

Качество ранжирования и многоязычное поведение

Таблица 1. Рейтинг по полному разработанному инструменту.

Метод	**nDCG@5 [95% CI]**	**MRR**	**P@5**	**R@5**
TF-IDF	0.312 [0.190, 0.464]	0,439	0,236	0,295
БМ25	0.314 [0.191, 0.467]	0,444	0,236	0,295
MiniLM-int8	0,685 [0,613, 0,754]	0,789	0,475	0,594

Метод	nDCG@5 [95% CI]	MRR	P@5	R@5
E5-маленький	0,792 [0,714, 0,854]	0,934	0,508	0,635

E5 лидирует в макросравнении nDCG, но на рисунке 2 показано языковое взаимодействие. MiniLM получает оценку 0,871 на английском и 0,826 на русском языках, по сравнению с E5, у которого оценки составляют 0,849 и 0,774. Для казахского языка E5 получает оценку 0,754, а MiniLM - 0,358. Разница между E5 и MiniLM для казахского языка составляет 0,396, с интервалом тематической бутстрап-анализа [0,252, 0,561] и условным p-значением, скорректированным по Холму, равным 0,0059. Контрасты между английским и русским языками не выдерживают этой коррекции.

TF-IDF дает минимальный результат по всем системам, но тем не менее, выборка у него была ниже всего по nDCG. Возможно, при увеличение 95% CI на несколько пунктов результат мог бы быть лучше, но в таком случае разница в MRR не была бы объективной.

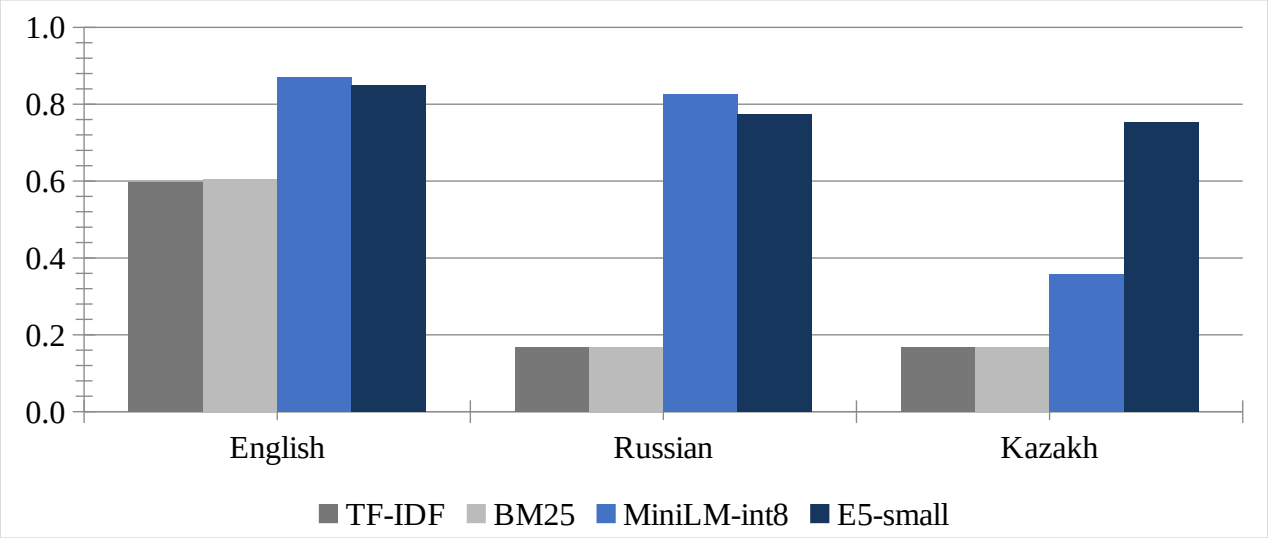


Рисунок 2. nDCG@5 по языку запроса по сравнению с английскими профилями.

Оба лексических метода выдают строки с нулевым значением для всех 48 запросов на русском и казахском языках. Их ненулевые показатели поиска по этим строкам обусловлены детерминированным упорядочиванием совпадений. Макроулучшение E5 по сравнению с TF-IDF составляет 0,480 [0,339, 0,603]; его условное скорректированное p равно 0,0107. Это различия, специфичные для конкретного инструмента, а не оценки институциональной выгоды.

Распределение ресурсов в условиях ограничений пропускной способности.

Таблица 2. Распределение мест по программе E5: 60 мест для 48 студентов.

Политика	Релевантность: среднее ± стандартное отклонение	Блокирующие пары	Индекс Джини	Утил. Джини
Случайный	0,116 ± 0,031	461.4	0,194	0,159
Жадный	0,713 ± 0,085	43.9	0,271	0.210
Гейл-Шэпли	0,639 ± 0,083	0.0	0,257	0,196
Максимум	0,720 ± 0,075	30.2	0,260	0,199

Точное присвоение максимального веса достигает релевантности 0,720, жадное - 0,713, а отложенное принятие - 0,639. Разница между парным точным и отложенным

принятием составляет 0,081, с интервалом бутстрапа сценария [0,061, 0,102]. Только отложенное принятие имеет нулевое количество блокирующих пар. Случайное присвоение имеет самый низкий коэффициент Джини нагрузки, несмотря на очень низкую релевантность. Таким образом, ни стабильность, ни равенство нагрузки сами по себе не являются достаточной мерой качества академического сопоставления.

Таблица 3. Чувствительность емкости E5 в 30 группах.

Емкость	Покрытие	Релевантность GS	Точная релевантность	GS Gini
48 мест	1.000	0.410	0,661	0.000
60 мест	1.000	0,639	0.720	0,257
36 мест	0,750	0,144	0,533	0,167

При наличии 48 мест все 24 руководителя должны принять по два студента, что обнуляет коэффициент Джини для каждой программы. При наличии 36 мест все места заполнены, коэффициент охвата составляет 0,750; коэффициент Джини нагрузки равен 0,167, а коэффициент Джини использования равен нулю для каждой программы. В условиях дефицита независимые предпочтения руководителей резко снижают значимость отложенного принятия; это проверка на наличие конфликта предпочтений.

Кодирование всех 96 текстов заняло 8,65 ± 0,50 с для MiniLM и 12,26 ± 0,22 с для E5 (среднее значение ± стандартное отклонение выборки; три повторных прохода). Повторные векторы были идентичны. Заключительные запуски отдельных процессов также точно воспроизвели все оценки, назначения и показатели результатов; затраченное время варьировалось.

Обсуждение

В рамках данной диагностики многоязычные представления улучшают средний рейтинг по сравнению с разработанной тематической рубрикой. Лексическое сравнение намеренно прямое и выявляет несоответствие языков. Оно не устанавливает превосходства над лексическим конвейером, который сначала переводит запросы, добавляет двуязычную терминологию или использует специфическую для языка морфологию. Такие меры контроля необходимы, прежде чем полностью объяснять наблюдаемый разрыв семантическим рассуждением.

Результаты, специфичные для разных языков, показывают, почему макроусреднение недостаточно. Две конфигурации встраивания меняют порядок описания в зависимости от языка. MiniLM остается конкурентоспособным на английском и русском языках, но значительно теряет в производительности на казахском. Таким образом, лучший макрорезультат E5 не является равномерным улучшением. Поскольку одна модель квантована, а другая - float32, эксперимент сравнивает развертываемые конфигурации, а не выделяет архитектуру или обучающие данные как причину. Для объяснения языкового разрыва необходимы сравнение с одинаковой точностью и проверка носителями языка.

Стабильное сопоставление и распределение по максимальному баллу реализуют разные стратегии. Отложенное принятие учитывает случайные приоритеты руководителей, даже если это снижает релевантность темы. Точное назначение принимает блокирующие пары, поскольку оно оптимизирует предпочтения студентов без ограничения стабильности. Эффект значительного дефицита возникает из-за полной приемлемости, ограниченного количества мест и независимо сформированных предпочтений руководителей. Замена этих предпочтений на полученные ответы или ранжирование, соответствующее релевантности, может существенно изменить компромисс.

Равенство рабочей нагрузки также требует интерпретации. Жесткие и дефицитные режимы вынуждают заполнять все доступные места, в результате чего статистика нагрузки становится следствием квот, а не свидетельством механизма справедливости. В условиях избытка случайное распределение более равномерно распределено в исходных данных, но гораздо менее значимо. Поэтому нормализованная по мощности загрузка указывается наряду

с исходным коэффициентом Джини. Ни один из индексов не измеряет количество часов надзора, сложность задач, дисциплинарные требования, справедливость в отношении защищенных групп или выгорание.

Полезный процесс принятия решений в университете должен, следовательно, отображать релевантность кандидатов и последствия распределения отдельно. Администраторам необходимо видеть незанятых студентов, потребление ресурсов, двусторонние конфликты и политику, которая привела к назначению. Прототип возвращает эти объекты и может принимать новые текстовые записи и явные списки предпочтений. Его результаты аудита поддерживают проверку; они не являются проверенными объяснениями измерений нейронного встраивания. Семантические объяснения и рабочий процесс институционального утверждения остаются задачами будущей реализации.

Эксперимент также устраняет проблему воспроизводимости, которая в противном случае осталась бы скрытой. Первоначальный независимый запуск выявил незначительные различия в BM25, вызванные неупорядоченным суммированием чисел с плавающей запятой; близкие совпадения приводили к изменению некоторых распределений. Сортировка уникальных поисковых запросов устранила этот источник. В окончательных полных повторных запусках используются разные начальные значения хешей Python, и сравниваются матрицы, метрики и назначения. Эта проверка важна, поскольку распределение может усиливать числовые различия, которые кажутся незначительными на этапе ранжирования.

Заключение

В этом исследовании теоретическая многоязычная архитектура сопоставления превращается в исполняемый, проверяемый эксперимент. E5 достигает наивысшего макрорейтинга по разработанному инструменту, в то время как языковой анализ показывает, что MiniLM остается более эффективным с точки зрения описания английского и русского языков. Отложенное принятие с учетом пропускной способности остается стабильным при заданных предпочтениях, но точное присвоение максимального веса обеспечивает большую релевантность разработанного инструмента. Равенство рабочей нагрузки может быть обеспечено режимом пропускной способности и не должно путаться с независимо продемонстрированной справедливостью.

На следующем этапе работы над диссертацией следует собрать согласованные исследовательские предложения и проверенные профили экспертов, получить анонимные двуязычные оценки релевантности от нескольких экспертов, выявить двусторонние предпочтения и оценить разногласия и согласованность оценок между экспертами. В качестве базового институционального корпуса следует сравнить лексический поиск с помощью перевода, многоязычные кодировщики с согласованной точностью и стабильную оптимизацию с дополнительными административными ограничениями. Выпущенный код и архивные матрицы обеспечат воспроизводимую отправную точку для этой проверки.

Доступность данных и кода. Сопутствующий исследовательский пакет содержит все созданные тексты и qrel, закрепленные манифесты моделей, исполняемый код, результаты для каждого запроса и каждого запуска, сохраненные оценки и эмбеддинги, точные инструкции по воспроизведению и аудит исследования. Он не содержит реальных студенческих записей. Веса предварительно обученной модели извлекаются отдельно из закрепленных официальных версий.

Список литературы:

1. Ван Л. Многоязычные текстовые эмбеддинги E5: технический отчёт [Электронный ресурс] / Л. Ван, Н. Ян, С. Хуан, Л. Ян, Р. Маджумдер, Ф. Вэй // arXiv. - 2024. - DOI: https://doi.org/10.48550/arXiv.2402.05672.

2. Гейл Д. Приём в колледжи и устойчивость брачных союзов / Д. Гейл, Л. С. Шепли // Американский математический ежемесячник. - 1962. - Т. 69, № 1. - С. 9-15. - DOI: https://doi.org/10.1080/00029890.1962.11989827.

3. Дрор Р. Руководство по проверке статистической значимости в обработке естественного языка / Р. Дрор, Г. Баумер, С. Шломов, Р. Райхарт // Материалы 56-й ежегодной конференции Ассоциации компьютерной лингвистики (Том 1: Полные статьи). - Ассоциация компьютерной лингвистики, 2018. - С. 1383-1392. - DOI: https://doi.org/10.18653/v1/P18-1128.

4. Мюннигхофф Н. MTEB: масштабный набор тестов для оценки текстовых эмбеддингов / Н. Мюннигхофф, Н. Тази, Л. Мань, Н. Раймерс // Материалы 17-й конференции Европейского отделения Ассоциации компьютерной лингвистики. - Ассоциация компьютерной лингвистики, 2023. - С. 2014-2037. - DOI: https://doi.org/10.18653/v1/2023.eacl-main.148.

5. Раймерс Н. Sentence-BERT: векторные представления предложений с использованием сиамских BERT-сетей / Н. Раймерс, И. Гуревич // Материалы конференции по эмпирическим методам обработки естественного языка и 9-й Международной объединённой конференции по обработке естественного языка (EMNLP-IJCNLP). - Ассоциация компьютерной лингвистики, 2019. - С. 3982-3992. - DOI: https://doi.org/10.18653/v1/D19-1410.

6. Раймерс Н. Создание многоязычных векторных представлений предложений на основе одноязычных моделей с использованием дистилляции знаний / Н. Раймерс, И. Гуревич // Материалы конференции по эмпирическим методам обработки естественного языка (EMNLP). - Ассоциация компьютерной лингвистики, 2020. - С. 4512-4525. - DOI: https://doi.org/10.18653/v1/2020.emnlp-main.365.

7. Рамотсиси Дж. Оптимизационная модель распределения студентов между руководителями проектов на примере инженерного факультета / Дж. Рамотсиси, М. Кгомотсо, Л. Себони // Журнал оптимизации. - 2022. - Т. 2022. - Статья 9415210. - DOI: https://doi.org/10.1155/2022/9415210.

8. Робертсон С. Вероятностная модель релевантности: BM25 и дальнейшее развитие / С. Робертсон, Х. Сарагоса // Основы и тенденции в информационном поиске. - 2009. - Т. 3, № 4. - С. 333-389. - DOI: https://doi.org/10.1561/1500000019.

9. Санчес-Ангикс В. Подход к распределению студентов между научными руководителями, близкий к оптимальному по Парето, с учётом двусторонних предпочтений и балансировки нагрузки / В. Санчес-Ангикс, Р. Чалумури, Р. Айдоган, В. Хулиан // Прикладные мягкие вычисления. - 2019. - Т. 76. - С. 1-15. - DOI: https://doi.org/10.1016/j.asoc.2018.11.049.

10. Серикбаева С. Тонкая настройка многоязычных трансформеров предложений для сопоставления навыков и понятий в условиях ограниченных языковых ресурсов на казахском, русском и английском языках / С. Серикбаева, М. Самбетбаева, В. Рамазанова, А. Еримбетова, З. Ламашева, З. Садирмекова, А. Батырханов, Е. Майлыбаев // Передовые исследования в области больших данных. - 2026. - Т. 9. - Статья 1900836. - DOI: https://doi.org/10.3389/fdata.2026.1900836.

11. Ярвелин К. Оценка методов информационного поиска на основе накопленного выигрыша / К. Ярвелин, Я. Кекяляйнен // Труды ACM по информационным системам. - 2002. - Т. 20, № 4. - С. 422-446. - DOI: https://doi.org/10.1145/582415.582418.