

Т.К. Буддыбаев  
А.О. Пилипенко  
И.Г. Акоева

АО «Информационно-аналитический центр» МОН РК, Нур-Султан, Казахстан  
(E-mail: i.akoyeva@gmail.com)

## Методика и мультиагентный аналитический инструментарий для разметки новостных текстов

**Аннотация.** В работе описан подход к созданию веб-ресурса по применению мультиагентного подхода для работы с размеченными корпусами новостных текстов. Подобная разметка позволит экспериментально подтвердить эффективность работы алгоритмов по распознаванию признаков в новостных текстах. Основными задачами исследования являются: 1) разработка онлайн модуля для разметки добровольцами новостных текстов; 2) использование полученных результатов разметки исследователями при работе с текстовыми корпусами; 3) сравнение результатов мультиагентной (добровольцы, эксперты, алгоритмы) разметки; 4) предоставление в свободный доступ размеченного корпуса новостных текстов по 3 классам: тональность (позитивная/нейтральная/негативная), объективность/субъективность, манипулятивность/неманипулятивность. Реализация инструментов разметки основывается на мультиагентном подходе, при котором учитываются 6 агентов формирования новостного контента, а также их взаимодействие в формировании повестки дня и в формировании культуры потребления новостей в целом.

Разработанный веб-ресурс, помимо основного функционала для разметки, позволяет осуществлять мониторинг результатов разметки с применением аналитических инструментов визуализации. Основным результатом исследования является получение размеченного корпуса новостных текстов различными агентами для разработки и тестирования алгоритмов распознавания информативных признаков в новостных текстах.

**Ключевые слова:** мультиагентный подход, СМИ, тональность, объективность, манипулятивность.

DOI: <https://doi.org/10.32523/2616-7174-2020-133-4-77-84>

Поступила: 08.09.20 / Доработана: 12.10.20 / Допущена к опубликованию: 20.12.20

Разработка информационных систем для анализа и мониторинга медиаконтента приобретает все большую актуальность в свете распространения фейковых и подозрительных публикаций.

Согласно глобальному отчету «Барометр доверия Эдельмана» 76% респондентов бес-

покоят использование ложной информации или фейковых новостей в качестве оружия [1]. 57% опрошенных, утверждают, что СМИ, которые они используют, содержат недостоверную информацию. При этом самые низкие показатели доверия к распространяемой новостной информации показывают

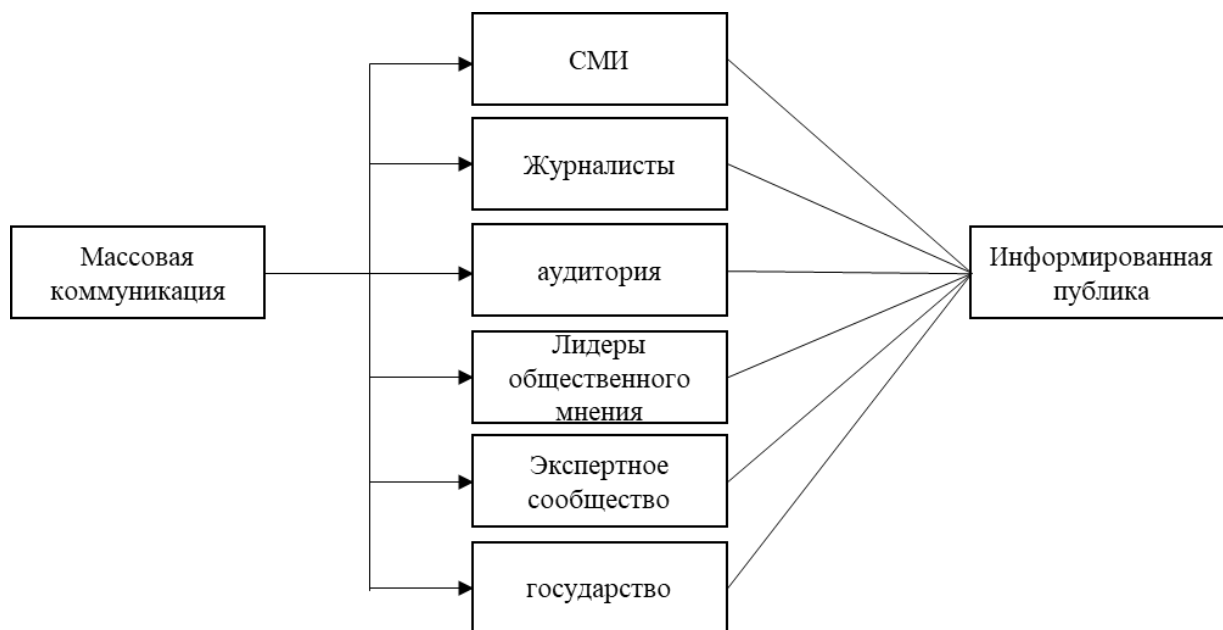


Рисунок 1. Типы агентов в теории массовой коммуникации

социальные сети. Наблюдается и глобальное недоверие к лидерам общества для решения проблем в стране. Так, 66% респондентов разделяют мнение, равное утверждению: «Я не уверен, что наши нынешние лидеры смогут успешно решить проблемы нашей страны». В свете данных тенденций, связанных с глобальным кризисом доверия к СМИ, учитывая повсеместное распространение фейковых и недостоверных новостей, а также падение доверия к институтам госуправления, разработка алгоритмов распознавания подозрительных публикаций является своевременной. Влияние СМИ на настроение общественности является также важной предпосылкой для основания проведения исследований и разработок в данной области.

В данной статье описано создание методики и веб-ресурса по применению мультиагентного подхода для создания размеченных корпусов новостных текстов с целью выработки алгоритмов распознавания информативных признаков влияния новостных текстов на социум.

**Мультиагентный подход к разметке корпуса.** Поскольку массовая коммуникация представляет собой сложный интерактивный процесс, предусматривающий активное взаимодействие различных агентов, участвующих в процессе, авторы исследования предполагают применение мультиагентного подхода к разметке корпуса для учета мнения и соблюдения паритета всех стейкхолдеров процесса массовой коммуникации. Мультиагентный подход предполагает, что при реализации определенных задач, в данном случае формирования экспертной разметки и классификации корпуса, привлекаются несколько агентов одновременно и их мнения учитываются при принятии решений или совершении окончательного выбора в рамках этих задач.

В теории массовой коммуникации и распространении новостей традиционно выделяют 6 типов агентов (Рисунок 1).

Каждый из таких агентов одновременно является актором. Однако в экосистеме массовой коммуникации ограничения на диапазон действий, определяемые функциями

**Финансирование:** Данная статья подготовлена в рамках реализации ПЦФ № BR05236839 по теме «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана».

внутри системы, для разных агентов различны. СМИ и частично государство формируют информационное поле и обладают наибольшей влиятельностью в части реализации этой функции. Теория массовой коммуникации оперирует такими понятиями, как создание инфоповодов, формирование повестки дня, теория шума и др. Государство одновременно участвует в создании информационного пространства и является зависимым от реакции аудитории. Аудитория выступает в качестве реципиента и распространителя новостной информации, а лидеры общественного мнения помимо этого выполняют функцию ее интерпретации, систематизации, уточнения. Функционал экспертного сообщества совпадает с функционалом лидеров общественного мнения, и в теории эксперт может являться лидером. Однако лидеры общественного мнения, даже при условии владения информацией в большем объеме нежели обычный представитель аудитории, не всегда демонстрируют надлежащий уровень экспертных знаний.

Вместе с тем в связи со стремительным развитием технологий и информационных каналов происходит трансформация функционала всех агентов внутри процесса массовой коммуникации. Представители аудитории часто становятся источниками новостной информации, формируют повестку дня, а СМИ и социальные сети выступают больше в роли ее распространителей.

Для разработки алгоритмов распознавания признаков в новостных текстах, основанных на методах машинного обучения с учителем, в первую очередь необходимо создание коллекции размеченных текстов. Следует отметить, что в Казахстане исследования в области корпусной лингвистики или разметки корпусов мало представлены [2].

Разметка и методологические аспекты формирования корпусов определяются практическим назначением корпуса. Исследователями применяются два подхода к разметке текстов: автоматизированная (массовая) или экспертная.

Библиотеки современных языков программирования включают многочисленные

инструменты автоматизированной разметки корпусов: корпусная статистика (количество слов, знаков, знаков препинания, предложений), морфологическая, частеречная, синтаксическая, семантическая, филологическая, библиографическая, тематическая, хронологическая, стилистическая и др.

Разметка корпусов, предусматривающая оценочные суждения человека, как правило, называется экспертной разметкой. В качестве экспертов к разметке корпусов привлекаются люди, чаще всего специалисты в определенной области, в соответствии с задачами разметки или исследования, в рамках которого происходит формирование корпусов. Даже при автоматизированной разметке корпусов финальный этап предусматривает валидацию результатов автоматизированной разметки человеком.

Несмотря на то, что автоматизация разметки является оптимальным решением для анализа больших массивов текстовых данных, при ее использовании сохраняется определенный процент погрешности (около 10-15%). Разметка, предусматривающая оценку прагматических аспектов текста и оценочных суждений о содержании текста, на современном этапе развития корпусной лингвистики не может обходиться программными и технологическими решениями и требует человеческого вмешательства. Как отмечают О.Н. Ляшевская и соавторы, на данный момент «не существует компьютерных программ, которые были бы способны заменить человека на этом направлении и обеспечить должный уровень адекватности» [3]. Примером исследования, показывающего несостоятельность автоматизированной разметки в определенных случаях, выступает исследование Й. Уилкса и соавторов по определению значения слова-омонима «bank», когда кластеризация слова на основе совместной встречаемости с другими словами позволила достичь точности автоматического распознавания 90%, но на уровне частных слов составила 53% [4]. И хотя проблема омонимов нашла приемлемое решение автоматизированного определения в целом, но в более сложных контекстуальных

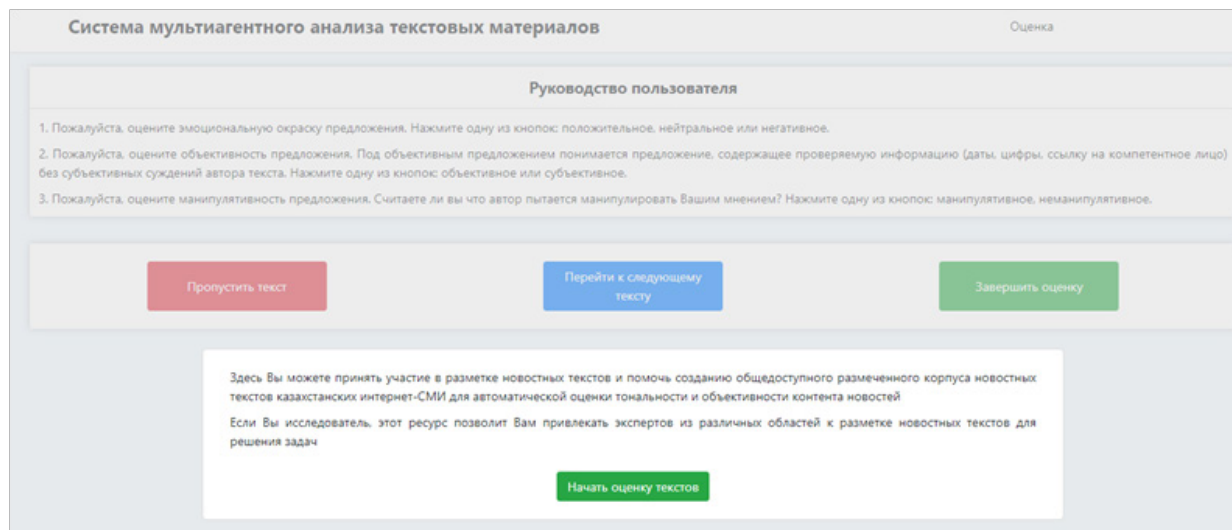


Рисунок 2. Система мультиагентного анализа текстовых материалов

задачах проблему оценки значений статистическими методами решить не удастся.

Учитывая, что целью нашего исследования является получение размеченного корпуса с применением мультиагентного подхода к оценке текстов, планируется привлечение добровольцев из различных областей, которые в свою очередь могут выступать в качестве экспертов в своей области.

Подобный подход применяется лабораторией интернет-исследований Высшей школы экономики города Санкт-Петербург. Исследователи создали краудсорсинговый веб-ресурс Linis Crowd для создания лингвистических инструментов, на основе массовой разметки добровольцами слов и текстов по тональности [5]. Результаты разметки используются для автоматической оценки тональности интернет-контента [6]. Существуют и другие краудсорсинговые платформы для разметки данных, такие как Amazon Mechanical Turk (MTurk), CrowdFlower, Яндекс.Толока. В целом краудсорсинг широко применяется при решении задач, не поддающихся автоматическим вычислениям и требующих человеческих усилий [7].

**Разработка веб-ресурса по разметке корпуса новостных текстов.** Для получения размеченного текста в рамках проекта был разработан веб-ресурс. В качестве разметки были определены 3 признака: (1) тональность

(позитивная/нейтральная/негативная), (2) объективность/субъективность, (3) манипулятивность/неманипулятивность (рисунок 2).

Суть разметки заключается в следующем. Для разметки добровольцами по 3 признакам необходимо прочитать представленные новостные тексты и ответить на следующие вопросы:

1. *Пожалуйста, оцените эмоциональную окраску предложения. Нажмите одну из кнопок: положительное, нейтральное или негативное.*

2. *Пожалуйста, оцените объективность предложения. Под объективным предложением понимается предложение, содержащее проверяемую информацию (даты, цифры, ссылку на компетентное лицо) без субъективных суждений автора текста. Нажмите одну из кнопок: объективное или субъективное.*

3. *Пожалуйста, оцените манипулятивность предложения. Считаете ли Вы, что автор пытается манипулировать Вашим мнением? Нажмите одну из кнопок: манипулятивное, неманипулятивное.*

При этом, если у оценщика возникнет трудность в разметке, он может пропустить данный текст (Рисунок 3).

Отображаемые публикации для разметки на веб-ресурсе генерируются автоматически – методом случайной выборки. Всего для разметки предлагаются публикации из корпуса объемом более 1,5 млн. публикаций, из ко-

**Система мультиагентного анализа текстовых материалов** Оценка

**Руководство пользователя**

- Пожалуйста, оцените эмоциональную окраску предложения. Нажмите одну из кнопок: положительное, нейтральное или негативное.
- Пожалуйста, оцените объективность предложения. Под объективным предложением понимается предложение, содержащее проверяемую информацию (даты, цифры, ссылку на компетентное лицо) без субъективных суждений автора текста. Нажмите одну из кнопок: объективное или субъективное.
- Пожалуйста, оцените манипулятивность предложения. Считаете ли вы что автор пытается манипулировать Вашим мнением? Нажмите одну из кнопок: манипулятивное, неманипулятивное.

Текст для разметки	Разметка
<p>Усть-Каменогорск, 21 апреля. "Казахстан Сегодня" - Лесохозяйственные учреждения ВКО не готовы к пожароопасному периоду. Об этом корреспонденту агентства сообщили в пресс-службе департамента по чрезвычайным ситуациям (ДЧС) ВКО. "В преддверии пожароопасного сезона в Восточном Казахстане проведены проверки готовности лесохозяйственных учреждений к тушению лесных пожаров и степени выполнения мероприятий по их предупреждению. Проверками были охвачены государственный природный резерват "Семей орманы" и Катон-Карагайский государственный национальный природный парк", - сообщили в ДЧС. По данным ДЧС, проверки показали, что, несмотря на проводимую работу по подготовке к пожароопасному периоду, имеется ряд проблемных вопросов, требующих немедленного выполнения. "В первую очередь это касается вопросов обучения и аттестации команд пожарно-химической станции, имеющихся технических неисправностей пожарной техники, недостатка средств пожаротушения. Не на должном уровне проводится работа с владельцами автодорог и линий электропередачи, проходящих по лесным массивам, по выполнению ими правил пожарной безопасности в лесах", - пояснили в ДЧС. По результатам проверок комиссией составлены предписания по устранению выявленных недостатков в кратчайшие сроки. Больше важных новостей в Telegram-канале «zakon.kz». Подписывайся!</p>	<p>Тональность <input type="radio"/> Позитивная <input type="radio"/> Нейтральная <input type="radio"/> Негативная</p> <p>Объективность <input type="radio"/> Объективная <input type="radio"/> Субъективная</p> <p>Манипулятивность <input type="radio"/> Манипулятивное <input type="radio"/> Неманипулятивное</p>

Пропустить текст
Перейти к следующему тексту
Завершить оценку

Рисунок 3. Разметка текстов на веб-ресурсе.

торых на первоначальном этапе апробацию прошли 500 текстов. Приоритетом для разметки выступают публикации с меньшим количеством полученных оценок.

Для мониторинга процесса заполняемости был разработан интерфейс формирования отчетности с применением аналитических инструментов визуализации (Рисунок 4).

После получения размеченного корпуса новостных тестов планируется проведение экспериментального подтверждения разработанных алгоритмов тональности, объективности и манипулятивности новостных текстов. Данные алгоритмы описаны в статьях [8, 9, 10]. Для анализа экспериментального подтверждения результатов массовой размет-

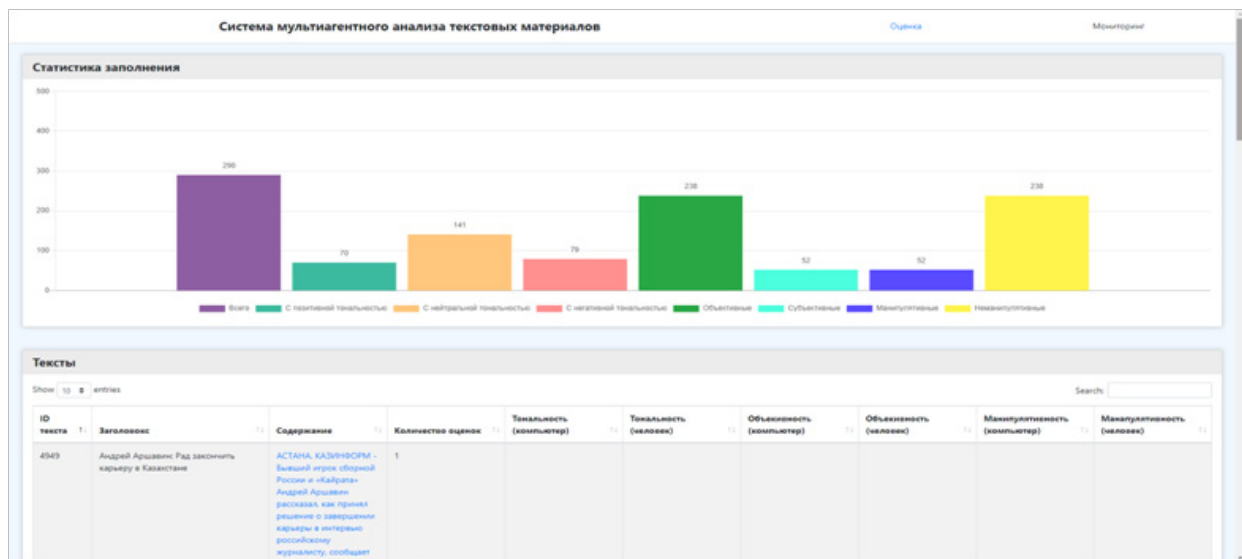


Рисунок 4. Модуль мониторинга результатов разметки и автоматического формирования отчетности

ки и алгоритмов возможно использовать параметры валидации Precision, Recall, f1-score. Для оценки согласия кодировки между добровольцами возможно использовать метрики Krippendorff's alpha coefficient или каппа Флейса.

**Заключение и выводы.** Таким образом, по результатам проведения комплекса работ был разработан веб-ресурс на основе мультиагентного подхода для создания размеченных корпусов новостных текстов. Разработанные веб-ресурс включают следующие функции: 1) по разметке новостных текстов по различной тематике по трем классам: тональность (позитивная/нейтральная/негативная), объективность/субъективность, манипулятивность/неманипулятивность; 2) мониторинг результатов разметки с применением аналитических инструментов визуализации. В рамках исследования данный ресурс решает задачи по 1) разработке онлайн модуля для разметки добровольцами новостных текстов; 2) предлагает функционал по использованию полученных результатов разметки исследователями при работе с текстовыми корпусами; 3) сравнению результатов мультиагентной (добровольцы, эксперты, алгоритмы) разметки; 4) предоставлению в свободный доступ размеченного корпуса новостных тестов по 3 классам.

Подобная разметка позволит создать размеченный корпус новостных текстов по 3 при-

знакам для увеличения размера обучающей выборки. Это позволит улучшить точность и полноту большинства прикладных решений, основанных на методах машинного обучения с учителем. Также результаты разметки позволяют экспериментально подтвердить эффективность работы алгоритмов распознавания признаков в новостных текстах.

В рамках следующего этапа развития системы по применению мультиагентного подхода к разметке текстов планируется имплементация алгоритмов сравнительного анализа компьютерной и пользовательской разметок по каждому из признаков. По отобранному ранее тексту будет проведена разметка с применением разработанных алгоритмов распознавания тональности, объективности и манипулятивности. Далее полученная алгоритмическим путем разметка будет сопоставлена с агрегированными пользовательскими разметками. В результате на веб-ресурсе будут представлены визуализация в виде графиков, отражающих коэффициенты соответствия между алгоритмическим и пользовательскими подходами в анализе, а также сводная таблица по проработанным текстам.

Основным результатом проекта является получение размеченного корпуса новостных текстов различными агентами для разработки и тестирования алгоритмов распознавания информативных признаков в новостных текстах.

### Список литературы

1. Edelman Trust Barometer 2018 Global study. [Электрон. ресурс] – URL: [http://cms.edelman.com/sites/default/files/201802/2018\\_Edelman\\_Trust\\_Barometer\\_Global\\_Report\\_FEB.pdf](http://cms.edelman.com/sites/default/files/201802/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf) (Дата обращения: 10.06.2020)
2. Оспанова У.А., Баймаханбетов М.А., Акоева И.Г., Буддыбаев Т.К., Атанаева М.К. Особенности формирования корпуса текстов казахстанского сегмента электронных новостей // Современные информационные технологии и ИТ-образование. - 2020. - № 1(16). – С. 46-58.
3. Ляшевская О.Н., Плунгян В.А., Сичинава Д.В. О морфологическом стандарте Национального корпуса русского языка // Научно-техническая информация. Серия 2: Информационные процессы и системы. – 2005. - № 6. –С. 111-135.
4. Wilks Y., Fass D., Guo C.-M., McDonald J.E., Plate T. and Slator B.M. Providing Machine Tractable Dictionary Tools // Machine Translation. –1990. – Vol. 5, -№2. – P. 99-154.
5. Алексеева С.В., Кольцов С.Н., Кольцова О.Ю. Linis-crowd.org: лексический ресурс для анализа тональности социально-политических текстов на русском языке // Компьютерная лингвистика и вычислительные онтологии. – Санкт-Петербург, 2015. – С. 25-32.

6. Koltsova O.Y., Alexeeva S.V., Kolcov S.N. An Opinion Word Lexicon and a Training Dataset for Russian Sentiment Analysis of Social Media // Компьютерная лингвистика и интеллектуальные технологии. – 2016. – С. 277–287.
7. Гилязов Р.А., Турдаков Д.Ю. Активное обучение и краудсорсинг: обзор методов оптимизации разметки данных // Труды ИСП РАН. – 2018. – №3(2). – С. 215-250.
8. Атанаева М.К., Булдыбаев Т.К., Оспанова У.А., Акоева И.Г., Нурумов К.С., Баймаханбетов М.А. Определение тональности и объективности новостных текстов словарным подходом // Научный аспект. – 2019. – №3(3). – С. 296-309.
9. Шакенова М.Т., Ташимханова Д.С., Оспанова У.А., Булдыбаев Т.К. Лексико-грамматические маркеры манипулятивности массмедийного дискурса (на материале корпуса русскоязычных текстов СМИ в Республике Казахстан) // Научный диалог. – 2020. – №2. – С. 141-160.
10. Кенжебалина Г.Н., Шаикова Г.К., Шакенова М.Т., Акоева И.Г. Распознавание манипулятивного текста: структура и доминантные языковые параметры // Научный диалог. – 2020. – №7. – С. 105-125.

Т.К. Булдыбаев, А.О. Пилипенко, И.Г. Акоева

ҚР БҒМ «Ақпараттық-талдау орталығы» АҚ, Нұр-Сұлтан, Қазақстан

### Жаңалықтар мәтіндерін айқындау әдістері мен мультиагентті сараптамалық құралдар

**Аңдатпа.** Мақалада жаңалықтар мәтіндерінің белгіленген корпустарымен жұмыс істеу үшін мультиагенттік тәсілді қолдану бойынша веб-ресурсты құру тәсілі сипатталған. Мұндай белгілеу жаңалықтар мәтіндеріндегі белгілерді тану алгоритмдерінің тиімділігін эксперименталды түрде растауға мүмкіндік береді. Зерттеудің негізгі міндеттері: еріктілердің жаңалық мәтіндерін белгілеуі үшін онлайн модуль әзірлеу; мәтіндік корпустармен жұмыс істеу кезінде зерттеушілермен алынған белгілеу нәтижелерін пайдалану; мультиагентті (еріктілер, сарапшылар, алгоритмдер) белгілеу нәтижелерін салыстыру; 3 сынып бойынша белгіленген жаңалықтар мәтіндерінің корпусына еркін қолжеткізуге мүмкіндік беру: үндестілік (оң/бейтарап/теріс), объективтілік/субъективтілік, манипулятивтілік/манипулятивтілік емес. Белгілеу құралдарын іске асыру, жаңалықтар контентін қалыптастыруда 6 агент ескерілетін, мультиагенттік тәсілге негізделген. Сондай-ақ, олардың күн тәртібін және жалпы жаңалықтарды тұтыну мәдениетін қалыптастырудағы өзара әрекеті қарастырылған.

Әзірленген веб-ресурс, белгілеуге арналған негізгі функционалдан басқа, визуализацияның аналитикалық құралдарын қолдана отырып, белгілеу нәтижелерін бақылауға мүмкіндік береді. Зерттеудің негізгі нәтижелері: (2) жаңалықтар мәтіндеріндегі ақпараттық белгілерді тану алгоритмдерін әзірлеу және тестілеу үшін (1) әртүрлі агенттердің ақпараттық мәтіндердің белгіленген корпусын алуы.

**Түйін сөздер:** мультиагенттік тәсіл, БАҚ, үндестілік, объективтілік, манипулятивтілік.

Т.К. Buldybayev, A.O. Pilipenko, I.G. Akoyeva

JSC «Information and Analytical Center» MES RK, Nur-Sultan, Kazakhstan

### Methodology and multi-agent analytical tools for marking up news texts

**Abstract.** The article describes an approach to creating a web resource for the application of a multi-agent model for working with marked-up corporate news texts. Such mark-up will allow to experimentally confirming the effectiveness of algorithms for feature recognition in news texts. The main objectives of the research are (1) development of an online module for marking-up news texts by volunteers; (2) the use of the obtained mark-up results by researchers when working with text corpora; (3) comparison of the results of multi-agent (volunteers, experts, algorithms) mark-up; (4) providing free access to a marked-up corpus of news texts in 3 classes: sentiment (positive/neutral/negative), objectivity/subjectivity, manipulative/non-manipulative. The implementation of markup tools is based on a multi-agent approach, which takes into account 6 agents for generating news content. And also, their interaction in shaping the agenda and, in general, shaping the culture of news consumption.

The developed web resource, along with the main functionality for mark-up, allows monitoring the mark-up results using analytical visualization tools. The main results of the study are: (1) obtaining a marked-up corpus

of news texts by various agents for (2) developing and testing algorithms for recognizing informative features in news texts.

**Keywords:** multi-agent approach, mass media, tonality, objectivity, manipulateness.

## References

1. Edelman Trust Barometer 2018 Global study [Electronic resource] Available at: [http://cms.edelman.com/sites/default/files/201802/2018\\_Edelman\\_Trust\\_Barometer\\_Global\\_Report\\_FEB.pdf](http://cms.edelman.com/sites/default/files/201802/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf) (Accessed: 10.06.2020)
2. Ospanova U.A., Baimahanbetov M.A., Akoyeva I.G., Buldybayev T.K., Atanayeva M.K. Osobennosti formirovaniya korpusa tekstov kazahstanskogo segmenta elektronnykh novostei [Features of the formation of the corpus of texts of the Kazakhstan segment of electronic news], *Sovremennye informacionnye tekhnologii i IT-obrazovanie* [Modern information technologies and IT education]. 2020. № 1(16). P. 46-58.
3. Lyashevskaya O.N., Plungyan V.A., Sichinava D.V. O morfologicheskom standarte Nacional'nogo korpusa russkogo yazyka [About the morphological standard of the National Corpus of the Russian language], *Nauchno-tekhnicheskaya informatsiya. Seriya 2: Informacionnye processy i sistemy* [Scientific and technical information. Series 2: Information processes and systems]. 2005. № 6. P. 111-135. [in Russian]
4. Wilks Y., Fass D., Guo C.-M., McDonald J.E., Plate T. and Slator B.M. Providing Machine Tractable Dictionary Tools, *Machine Translation*. 1990. Vol. 5. №2. P. 99-154.
5. Alekseeva S.V., Kolcov S.N., Kolcova O.Yu. Linis-crowd.org: leksicheskii resurs dlya analiza tonalnosti socialno-politicheskikh tekstov na russkom yazyke [Linis-crowd.org: Lexical resource for analyzing the tonality of socio-political texts in Russian], *Kompyuternaya lingvistika i vychislitelnye ontologii* [Computational linguistics and computational ontologies]. Saint Petersburg. 2015. P. 25-32. [in Russian]
6. Koltsova O.Y., Alexeeva S.V., Kolcov S.N. [An Opinion Word Lexicon and a Training Dataset for Russian Sentiment Analysis of Social Media], *Kompyuternaya lingvistika i intellektualnye tekhnologii* [Computational linguistics and intelligent technologies]. 2016. P. 277-287.
7. Gilyazev R.A., Turdakov D.YU. Aktivnoe obuchenie i kraudsorsing: obzor metodov optimizatsii razmetki dannykh [Active learning and crowdsourcing: an overview of data markup optimization methods], *Trudy ISP RAN* [Proceedings of the ISP RAS]. 2018. №3(2). P. 215-250. [in Russian]
8. Atanayeva M.K., Buldybayev T.K., Ospanova U.A., Akoyeva I.G., Nurumov K.S., Baimahanbetov M.A. Opredelenie tonalnosti i obektivnosti novostnykh tekstov slovarnym podhodom [Determining the tonality and objectivity of news texts using a dictionary approach], *Nauchnyy aspekt* [The scientific aspect]. 2019. №3(3). P. 296-309. [in Russian]
9. Shakenova M.T., Tashimhanova D.S., Ospanova U.A., Buldybayev T.K. Leksiko-grammaticheskie markery manipulyativnosti massmedijnogo diskursa (na materiale korpusa russkoyazychnykh tekstov SMI v Respublike Kazahstan) [Lexical and grammatical markers of manipulateness of mass media discourse (based on the corpus of Russian-language media texts in the Republic of Kazakhstan)], *Nauchnyy dialog* [Scientific dialogue]. 2020. №2. P. 141-160. [in Russian]
10. Kenzhebalina G.N., Shaikova G.K., Shakenova M.T., Akoyeva I.G. Raspoznavanie manipulyativnogo teksta: struktura i dominantnye yazykovye parametry [Manipulative text recognition: structure and dominant language parameters], *Nauchnyy dialog* [Scientific dialogue]. 2020. №7. P. 105-125. [in Russian]

## Сведения об авторах:

**Булдыбаев Тимур – основной автор**, директор Департамента прикладных исследований и разработок, АО «Информационно-аналитический центр», Нур-Султан, Казахстан.

**Пилипенко Александр** – инженер-программист Департамента прикладных исследований и разработок, АО «Информационно-аналитический центр», Нур-Султан, Казахстан.

**Акоева Инесса** – главный аналитик Департамента прикладных исследований и разработок, АО «Информационно-аналитический центр», Нур-Султан, Казахстан.

**Buldybaev Timur** – **main author**, Director of the Department of Applied Research and Development, «Information-Analytical Center» JSC, Nur-Sultan, Kazakhstan.

**Pilipenko Aleksandr** – Software Engineer of the Department of Applied Research and Development, «Information-Analytical Center» JSC, Nur-Sultan, Kazakhstan.

**Akoyeva Inessa** – Chief Analyst of the Department of Applied Research and Development, «Information-Analytical Center» JSC, Nur-Sultan, Kazakhstan.