

ОТ РАЗМЕЧЕННОГО КОРПУСА К ОНТОЛОГИИ: ФОРМАЛЬНАЯ МОДЕЛЬ МНОГОУРОВНЕВОЙ СЕМАНТИЧЕСКОЙ РАЗМЕТКИ ФЕЙКОВЫХ НОВОСТЕЙ В КАЗАХСКО-РУССКОМ МЕДИАПРОСТРАНСТВЕ (KazFakeCorpus/KazFakeOnto)

А. А. Некесова¹ [0009-0001-7642-0648], Е. А. Сидорова² [0000-0001-8731-3058],
Ж. Б. Ламашева³ [0000-0002-9535-2636], М. А. Самбетбаева⁴ [0000-0001-9358-1614],
А. С. Еримбетова⁵ [0000-0002-2013-1513], М. Ж. Калдарова⁶ [0000-0001-7494-9794],
А. А. Назымхан⁷ [0009-0007-4366-6709]

^{1, 3, 4, 6, 7}Международный научный комплекс «Астана», г. Астана, Казахстан

^{1, 3, 4, 7}Евразийский национальный университет имени Л. Н. Гумилева,
г. Астана, Казахстан

²Институт систем информатики им. А. П. Ершова СО РАН, г. Новосибирск,
Россия

⁵Институт информационных и вычислительных технологий КН МНВО РК,
г. Алматы, Казахстан

⁵Школа инженерии и информационных технологий, Университет МЕНА,
г. Алматы, Казахстан

⁶Международный университет Астана, г. Астана, Казахстан

¹nekessova_aa_1@enu.kz, ²lsidorova@iis.nsk.su, ³lamasheva_zhb@enu.kz,

⁴sambetbayeva_ma_1@enu.kz, ⁵aigerian8888@gmail.com,

⁶mira_kaldarova@aiu.edu.kz, ⁷nazymkhan_aa@enu.kz

Аннотация

Рассмотрена проблема формального представления знаний о дезинформации в малоресурсном двуязычном медиапространстве. Автоматическое обнаружение фейковых новостей традиционно сводится к бинарной классификации REAL/FAKE, которая не описывает внутреннюю структуру недостоверного сообщения: тип фейкового контента, примененную технику дезинформации, коммуникативное намерение автора, характеристики источника и доказательной базы.

Представлены два взаимосвязанных ресурса. Первый – это двуязычный корпус KazFakeCorpus: 4276 текстов на казахском и русском языках, сбалансированных по классам внутри каждого языка; класс REAL сформирован из официальных материалов портала Gov.kz, класс FAKE – из их контролируемых трансформаций. Разметку по многоуровневой семантической схеме выполнили в среде Label Studio два независимых эксперта. Второй ресурс – это онтология KazFakeOnto, переводящая схему разметки из плоского реляционного представления в формальную модель на языке OWL 2 DL: 38 классов, 26 объектных свойств, 22 свойства данных, 112 индивидов, 1155 утвержденных триплетов. Одиннадцать определяемых классов не утверждаются вручную, а выводятся ризонером из значений свойств. Непротиворечивость проверена с помощью HermiT, запросный слой апробирован на пятнадцати SPARQL- и тринадцати DL-запросах.

Показано, что онтологический слой обеспечивает воспроизводимую спецификацию схемы разметки, автоматический вывод содержательных категорий и единую формальную среду для постановки корпусно-лингвистических вопросов и анализа ошибок моделей машинного обучения. Обсуждены ограничения подхода и направления масштабирования онтологии на полный корпус и на естественно возникающую дезинформацию.

Ключевые слова: обнаружение фейковых новостей, онтология, OWL 2, SPARQL, логический вывод, двуязычный корпус, казахский язык, многоуровневая аннотация, техники дезинформации, объяснимый NLP, малоресурсные языки, KazFakeCorpus, KazFakeOnto.

ВВЕДЕНИЕ

Распространение цифровых медиа привело к росту объема недостоверных сообщений, и автоматическое обнаружение фейковых новостей закрепилось как самостоятельная задача обработки естественного языка [1]. Качество таких систем в решающей степени определяется качеством корпусных ресурсов, а подавляющее большинство доступных наборов данных построено для высокоресурсных языков [2, 3]. Для казахского языка проблема стоит особенно остро: при наличии корпусов общего назначения [4] ресурсы, пригодные для анализа дезинформации, практически отсутствуют. Дополнительную сложность создает

двуязычная медиасреда Казахстана, в которой казахский и русский языки используются параллельно.

Существенным ограничением большинства ресурсов является сама постановка задачи. Бинарная схема REAL/FAKE фиксирует итоговый вердикт, но не сохраняет тех сведений, ради которых вердикт и выносится: какого рода искажение допущено, какими риторическими средствами оно оформлено, на что рассчитывал автор, насколько надежен источник. Между тем именно эти сведения нужны и эксперту-фактчекеру, и разработчику объяснимой модели.

В работах [5–7] нами предложены многоуровневая схема семантической разметки фейковых новостей и построенный по ней двуязычный корпус KazFakeCorpus. Схема реализована в среде Label Studio в виде набора связанных сущностей и полей и представляет собой, по существу, ER-модель предметной области. Такая форма удобна для организации ручной разметки, однако семантика полей задается в ней неформально, корректность заполнения не контролируется автоматически, а агрегирование сведений по нескольким уровням требует отдельного программного кода.

Настоящая статья является расширенной версией доклада [7], представленного на конференции «Диалог-2026». По отношению к [5, 6] новыми являются онтологическая формализация схемы (разд. 4), запросный слой и результаты логического вывода (разд. 5). Онтология не заменяет корпус, а надстраивается над ним: она фиксирует смысл каждого уровня разметки в виде логических аксиом, позволяет ризонеру выводить принадлежность материала к содержательным категориям вместо ручного проставления меток и предоставляет общий декларативный язык запросов к разметке.

Вклад работы состоит в следующем: многоуровневая схема разметки формализована средствами OWL 2 DL; построена онтология KazFakeOnto, охватывающая уровни разметки, участников аннотирования, источники, доказательную базу, эвиденциальные маркеры, модели детекции и порождаемые ими объяснения; предложены одиннадцать определяемых классов, выводимых ризонером из значений свойств, а непротиворечивость модели подтверждена ризонером HermiT; апробирован запросный слой из пятнадцати SPARQL- и тринадцати DL-запросов.

1. БЛИЗКИЕ ПО ТЕМАТИКЕ ИССЛЕДОВАНИЯ

Корпусные ресурсы. Состояние области представлено в обзоре [1]: подавляющее большинство ресурсов построено для высокоресурсных языков и ограничено бинарной разметкой достоверности; социальный контекст и метаданные учитываются лишь в отдельных наборах данных – FakeNewsNet [8] (~23 тыс. материалов) и MCFEND [9] (~23.8 тыс.). Наиболее глубоко размечен немецкий корпус DeFaktS [10]: ~ 105 тыс. сообщений, десять признаков дезинформации одновременно со степенью достоверности. Для малоресурсных языков доступные корпуса невелики: 900 документов для урду [2], около 15 тыс. записей для курдского языка [3]. KazFakeCorpus (4276 текстов) уступает крупнейшим ресурсам по объему и отличается происхождением класса FAKE: материалы порождены контролируемой трансформацией верифицированных сообщений, а не собраны в естественной среде. Такой дизайн обеспечивает попарную сопоставимость классов, но ограничивает внешнюю валидность (разд. 6).

Многоуровневые схемы разметки. Дезинформация представляет собой не однородное явление, а совокупность форм и приемов [11], что мотивирует переход от бинарного вердикта к многомерным схемам. Наиболее развитой линией здесь является разметка манипулятивных приемов: в SemEval-2020 Task 11 [12] выделены восемнадцать пропагандистских техник с привязкой к текстовым фрагментам. Таксономия TAXODIS [13] систематизирует лингвистические признаки дезинформационного дискурса и опубликована в виде SKOS-словаря – первый известный нам случай формальной сериализации схемы разметки дезинформации. Существенно, однако, что SKOS фиксирует иерархию понятий и не поддерживает логического вывода: словарь задает допустимые значения меток, но не позволяет вычислять производные категории материала. Именно этот разрыв составляет предмет настоящей работы, которая не расширяет перечень меток [5, 6], а изменяет способ их представления.

Формальные модели и графы знаний. Словарь ClaimReview консорциума schema.org [14] стандартизирует публикацию результатов фактчекинга и служит основой для агрегирующих графов знаний: ClaimsKG [15] собирает проверенные утверждения с нормализованными вердиктами, CimpleKG [16] расширяет этот

подход до непрерывно обновляемого ресурса, объединяющего данные 77 факчекинговых организаций и свыше 200 тыс. документов на 26 языках. Эти ресурсы отвечают на вопрос, кто, когда и с каким вердиктом проверил утверждение, но не описывают внутреннюю структуру недостоверного сообщения – примененную технику, тип искажения, намерение автора.

Онтология Fandet [17] формализует средствами OWL контекстные признаки публикации в социальных медиа, но описывает преимущественно ее окружение – автора, платформу, распространение, а не семантику искажения, и не содержит конструкций логического вывода. Работа [18], выполненная в казахстанском контексте, близка к настоящей по материалу и инструментам, однако в ней моделируются акторы, целевые аудитории и эмоциональные оценки информационного воздействия, а онтология служит средством структурирования графа; в KazFakeOnto же содержательные категории материала выводятся ризонером из значений элементарных свойств.

Отметим разделение труда между OWL и языком ограничений SHACL: требования вида «поле обязательно заполнено» в OWL не проверяются вследствие допущения открытого мира, поэтому контроль полноты разметки требует SHACL, тогда как OWL 2 DL необходим для вывода производных категорий; в настоящей работе формализуется вторая задача. Архитектурная роль онтологии как слоя интеграции над разнородными ресурсами проработана в проектах цифровых научных экосистем, характерным примером которых служит экосистема OntoMath [19]; настоящая работа переносит эту идею – онтология над корпусом, а не вместо него – в область анализа дезинформации.

Таким образом, лакуна локализуется по трем признакам одновременно. Формальные ресурсы уровня связанных данных [14–16] ориентированы на метаданные проверки и не описывают внутреннюю структуру сообщения. Модели, описывающие эту структуру [17, 18], либо ограничены контекстом публикации, либо не доводят формализацию до логического вывода. Схемы многоуровневой разметки [5, 6, 10, 13] формализованы средствами, не поддерживающими вывода, и ни одна из них не охватывает процесс аннотирования, результаты моделей детекции и порождаемые ими объяснения в едином графе; для казахско-русского материала не существует ресурса ни одного из перечисленных типов.

2. KazFakeCorpus: ПОСТРОЕНИЕ КОРПУСА

Общая схема исследования приведена на рис. 1: верхний ряд соответствует построению корпуса, нижний – его онтологической формализации и запросному слою.



Рис. 1. Схема исследования: от сбора данных к онтологической модели и запросному слою.

Материалы класса REAL. Официальные новостные материалы собирались с портала gov.kz по тематике цифровизации, искусственного интеллекта и государственных инициатив. Ограничение корпуса одной предметной областью снижает тематическую вариативность и обеспечивает контролируемое сопоставление классов. Материалы прошли дополнительную проверку экспертами предметной области до начала разметки.

Материалы класса FAKE. Тексты класса FAKE создавались путем контролируемой трансформации материалов класса REAL с сохранением общей тематической структуры и внесением целенаправленных искажений, соответствующих заранее определенным техникам дезинформации. Для генерации использовалась модель GPT-5 в интерфейсе ChatGPT: на вход подавался исходный текст T и системная инструкция, предписывающая проанализировать его фактологическую структуру, выбрать релевантные значения категорий $FAKE_type$ и $DISINFORMATION_TECHNIQUE$ и выполнить преобразование с сохранением тематической связности; результатом являлся синтетический пример T' . Каждый порожденный текст проходил экспертную проверку; материалы с прямым копированием исходного текста, логическими противоречиями, явными языковыми

ошибками либо не соответствовавшие заданной технике исключались.

Такой дизайн минимизирует вариативность, не связанную с дезинформацией: каждый экземпляр класса FAKE порожден из соответствующего экземпляра класса REAL, поэтому наблюдаемые различия локализованы на уровне семантической манипуляции. Отношение порождения фиксируется свойствами `generatedBy` и `derivedFromSeed` (разд. 4), что позволяет ставить вопрос о том, какие признаки – дезинформационные или порождающие – усваивает обучаемая модель.

Предобработка и разбиение. Предобработка включала удаление дубликатов, фильтрацию шума, удаление URL-адресов и нормализацию казахского и русского текстов; дедупликация применялась внутри класса REAL, поскольку пары REAL/FAKE близки по содержанию по построению. Из того же обстоятельства следует требование к разбиению на обучающую и тестовую части: случайное разбиение по документам развело бы порождающий и порожденный тексты по разным подвыборкам и зависило бы оценки качества, поэтому разбиение выполняется по группам, объединяющим исходный материал и все порожденные из него тексты; это требование должно соблюдаться и в последующих работах, использующих корпус. Корпус сбалансирован по классам и языкам (табл. 1).

Табл. 1. Статистика корпуса KazFakeCorpus.

Язык	Класс	Число текстов	Среднее число слов	Доля, %
Казахский	REAL	1085	≈ 278	25.4
Казахский	FAKE	1085	≈ 330	25.4
Русский	REAL	1053	≈ 301	24.6
Русский	FAKE	1053	≈ 448	24.6
Всего	–	4276	≈ 339	100

Обращает на себя внимание систематическое различие средней длины текстов классов REAL и FAKE, особенно выраженное для русского языка (301 против 448 слов). Это потенциальный поверхностный признак, на который может опереться классификатор, поэтому при оценке моделей на корпусе целесообразно приводить базовое решение, использующее только длину текста; фиксация

этого признака – одно из оснований для включения свойства tokenCount в онтологическую модель.

Протокол аннотирования. Разметка выполнялась двумя независимыми экспертами по предопределенной многоуровневой схеме в среде Label Studio; схема уточнялась на пилотном этапе, охватившем 120 текстов. Помимо категориальных меток каждый аннотатор фиксировал уверенность в решении по пятибалльной шкале. Материалы с расхождением решений и с низкими значениями уверенности передавались на дополнительное рассмотрение; итоговая метка принималась по консенсусу экспертов. Сохраняемая в корпусе оценка уверенности переносится в онтологию как значение свойства annotatorConfidence и используется в определении класса HighConfidenceFake.

Согласованность оценивалась коэффициентом альфа Криппендорфа (Krippendorff's Alpha) [20] по разметке двух аннотаторов до этапа согласования на случайной подвыборке в 15 % корпуса. Получены следующие значения α : достоверность (бинарная номинальная переменная) – 0.88; тип фейкового контента – 0.84; техника дезинформации – 0.81; модальность – 0.79.

Значения α в диапазоне 0.79–0.88 свидетельствуют о высокой согласованности. Наиболее низкие значения получены для техники дезинформации и модальности: эти уровни требуют интерпретации текста и допускают приписывание одному материалу нескольких значений, поэтому коэффициент вычислялся для них в варианте для множественных ответов. Тем же продиктованы ограничения на свойства в онтологии: hasVeracityLabel и hasLanguage объявлены функциональными, тогда как usesTechnique, hasModality и hasFakeNewsType таких ограничений не имеют.

3. МНОГОУРОВНЕВАЯ СХЕМА СЕМАНТИЧЕСКОЙ РАЗМЕТКИ

Центральная сущность. Центральным элементом схемы является сущность NEWS – базовая единица описания новостного сообщения. Она агрегирует метаданные, необходимые для формальной идентификации и контекстуализации текста: уникальный идентификатор, дату разметки, итоговую метку достоверности REAL_or_FAKE, характеристики источника source_type, временные и географические параметры, целевую сущность TARGET_entity и оценку уверенности CONFIDENCE по шкале 1–5. В зависимости от значения REAL_or_FAKE сущность

NEWS связывается с блоком REAL_NEWS либо с блоком FAKE_NEWS.

Блоки REAL_NEWS и FAKE_NEWS. Блок REAL_NEWS содержит параметры, отражающие механизмы верификации: EVIDENCE и SOURCE_CREDIBILITY, что согласуется с подходами автоматизированного фактчекинга, где достоверность определяется проверяемостью утверждения и надежностью источника.

Блок FAKE_NEWS описывает внутреннюю структуру дезинформации и включает fake_type, DISINFORMATION_TECHNIQUE и AUTHOR_INTENT. Компонент fake_type опирается на устоявшиеся типологии недостоверного контента [11] и охватывает тринадцать подтипов, среди которых отметим сфабрикованный контент, вводящий в заблуждение контент и ложный контекст. Компонент DISINFORMATION_TECHNIQUE описывает манипулятивные стратегии – кликбейт, эмоциональное давление, преувеличение, преуменьшение, поиск виновного и ложную атрибуцию; в корпусе преобладают ложная атрибуция, кликбейт и эмоциональное давление, однако это распределение отражает состав инструкций, использованных при контролируемой трансформации, а не наблюдаемую частотность приемов. Включение AUTHOR_INTENT добавляет разметке прагматическое измерение, блок MODALITY фиксирует форму представления содержания (текст, изображение, видео, аудио) и делает схему расширяемой на мультимодальные постановки задачи. Модель включает также явные семантические отношения между сущностями, что предоставляет возможность графового представления корпуса.

Ограничения плоского представления. Реализация схемы в виде связанных таблиц разметки имеет три ограничения. Во-первых, семантика полей задается неформально: утверждение «материал класса FAKE не может иметь заполненное поле EVIDENCE» существует лишь в инструкции для аннотатора. Во-вторых, содержательные категории приходится вводить вручную: чтобы выделить, например, казахскоязычные фейки, порожденные языковой моделью и содержащие эвиденциальный маркер, требуется либо новое поле разметки, либо отдельный программный запрос. В-третьих, эталонная разметка, предсказания моделей, объяснения и метаданные корпуса хранятся в разных форматах, и их совместный анализ требует связующего кода.

Онтологическое представление устраняет второе и третье ограничения полностью: содержательные категории выводятся ризонером, а разнородные

сведения объединяются в едином графе, доступном для декларативных запросов. Первое ограничение снимается частично: аксиомы OWL позволяют обнаружить противоречащие друг другу утверждения, но не выявляют отсутствие обязательного значения.

4. KazFakeOnto: ОНТОЛОГИЧЕСКАЯ ФОРМАЛИЗАЦИЯ СХЕМЫ

Принципы проектирования. Онтология KazFakeOnto построена на языке OWL 2 DL [21] в среде Protégé; идентификатор онтологии – <http://www.semanticweb.org/kazfake>, префикс пространства имен – `kf:`. Соблюдались четыре принципа. Первый – это соответствие схеме разметки: каждый ее уровень отображается в конструкцию онтологии, что позволяет автоматически преобразовывать выгрузку корпуса в RDF-триплеты (целевая сущность и географические параметры сущности NEWS пока не формализованы). Второй – справочники представлены именованными индивидами классов-справочников, а не классами, что упрощает построение SPARQL-запросов. Третий – содержательные категории не проставляются вручную, а определяются аксиомами эквивалентности и вычисляются ризонером. Четвертый – это проверяемость: каждая конструкция сопровождается запросом, результат которого сопоставляется с ожидаемым.

Количественные характеристики модели: 1155 утвержденных триплетов (без учета выведенных), 38 именованных классов, 26 объектных и 22 свойства-значения данных, 112 индивидов. Наполнение включает демонстрационный фрагмент корпуса из 18 размеченных материалов (9 казахских и 9 русских, 12 FAKE и 6 REAL, из них 3 машинно-порожденные), а также источники, аннотаторов, объекты аннотации, эвиденциальные маркеры, подтверждения, порождающие модели, модели детекции с метриками, предсказания, объяснения и мета-данные корпуса. Фрагмент носит иллюстративный характер: он построен так, чтобы задействовать все конструкции модели, поэтому приводимые далее количественные результаты характеризуют состав фрагмента, а не распределения в корпусе.

Структура классов. Таксономия классов представлена на рис. 2. Классы-контейнеры описывают основные типы сущностей предметной области: `Corpus`, `NewsItem`, `Annotator`, `Annotation`, `Source`, `Evidence`, `DetectionModel`, `Prediction`, `Explanation`, `EvidentialMarker`, `Generator`. Класс `Explanation` разделен на подклассы

CommonsenseExplanation, TextualExplanation и EvidenceBasedExplanation. Класс EvidentialMarker разделен на EvidentialAffix, EvidentialParticle и ReportativeConstruction – это разбиение отражает грамматическую типологию средств выражения эвиденциальности [22], существенно различающихся в казахском и русском языках.

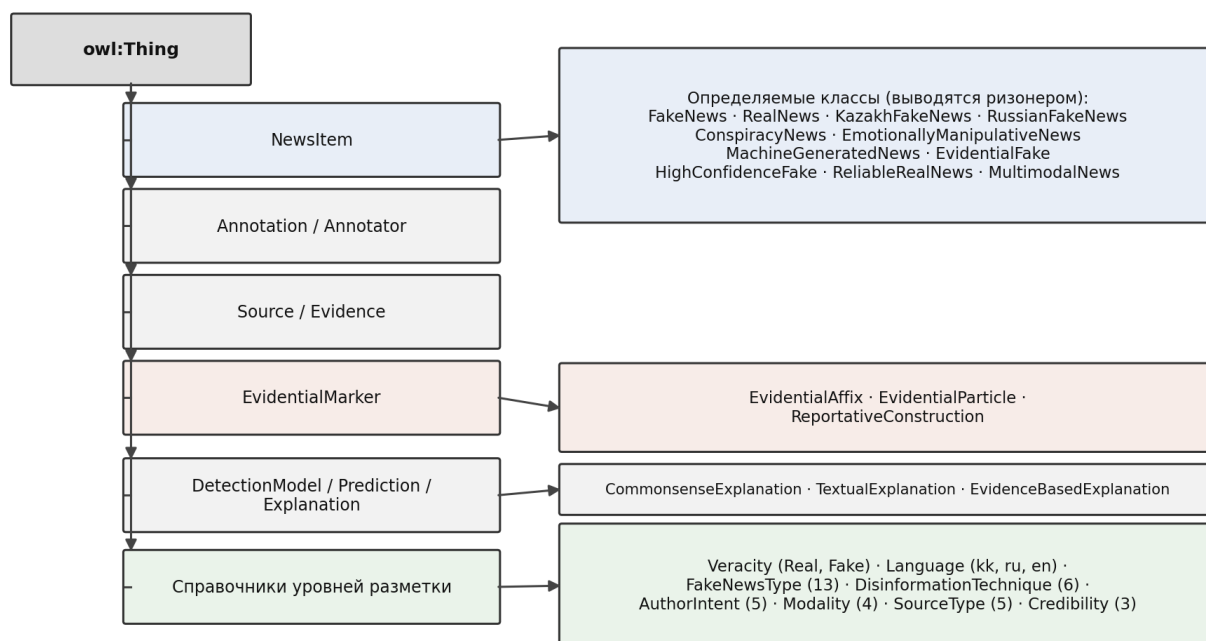


Рис. 2. Фрагмент таксономии классов KazFakeOnto.

Восемь классов-справочников отображают уровни разметки в конструкции онтологии:

- достоверность – Veracity и функциональное свойство hasVeracityLabel (2 значения);
- язык – Language и функциональное hasLanguage (3, одно зарезервировано); тип фейкового контента – FakeNewsType и hasFakeNewsType (13);
- техника дезинформации – DisinformationTechnique и usesTechnique (6); интенция автора – AuthorIntent и hasAuthorIntent (5);
- модальность – Modality и hasModality (4); тип источника – SourceType и hasSourceType (5);
- доверие к источнику – Credibility и функциональное hasCredibility (3).

Первые шесть свойств имеют доменом NewsItem, два последних – Source. Индивиды каждого справочника объявлены попарно различными аксиомой

owl:AllDifferent: в OWL не действует допущение уникальности имен, и без этого объявления ризонер не смог бы ни установить нарушение функциональности свойства, ни вывести принадлежность материала к классу, определенному через ограничение мощности. Значение MixedLanguage предназначено для материалов со смешением кодов; в текущей версии корпуса они отсутствуют, значение включено для расширяемости модели.

Свойства и аксиомы. Свойства сущности NewsItem образуют звездообразную структуру (рис. 3). Дататипные свойства фиксируют идентификатор newsId, заголовок title, дату newsDate, объем tokenCount, уверенность аннотатора annotatorConfidence и примечания notes.

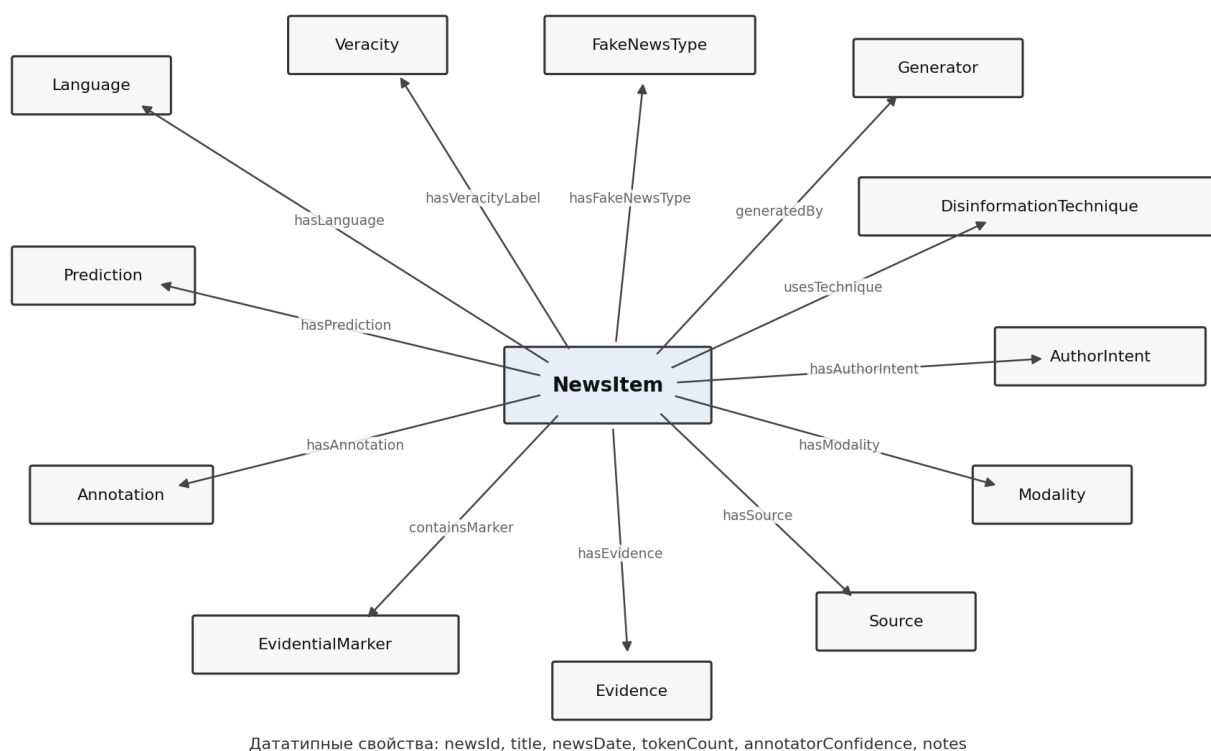


Рис. 3. Граф свойств сущности NewsItem.

Помимо доменов и областей значений заданы четыре типа содержательно значимых аксиом. Свойства hasVeracityLabel, hasLanguage, hasCredibility и madeBy объявлены функциональными: материал не может одновременно относиться к классам REAL и FAKE, иметь два языка разметки, а источник – два уровня доверия. Само по себе это не гарантирует непротиворечивости: не располагая сведениями о различии индивидов, ризонер отождествил бы kf:Fake и

kf:Real. Поэтому функциональность дополнена объявлением индивидов справочников попарно различными и аксиомой дизъюнктивности классов FakeNews и RealNews. Пара обратных свойств hasAnnotation/annotates позволяет обходить граф в обоих направлениях без дублирования утверждений.

Аксиома-цепочка SubObjectPropertyOf(ObjectPropertyChain(hasAnnotation madeBy) annotatedBy) связывает материал непосредственно с аннотатором. В OWL 2 цепочка свойств выражается только включением, но не эквивалентностью; вследствие этого свойство annotatedBy становится несоставным и не может использоваться в ограничениях мощности и объявляться функциональным, иначе онтология вышла бы за пределы OWL 2 DL. Свойство annotatedBy нигде не утверждается явно: все его значения выводятся ризонером, что демонстрирует запрос D10.

Свойства происхождения generatedBy и derivedFromSeed фиксируют дизайн контролируемой трансформации (разд. 2) и делают воспроизводимой проверку гипотезы о влиянии порождающей модели на обучаемый детектор.

Отметим ограничение текущей версии, существенное для интерпретации (разд. 5): уровни разметки заданы как свойства сущности NewsItem, а не объекта Annotation. Это соответствует хранению согласованного результата разметки, но не позволяет представить в графе две независимые разметки одного материала, поэтому пересчет коэффициента согласованности по графу невозможен.

Определяемые классы. Ключевым элементом, отличающим онтологию от структурированной таблицы, являются определяемые классы; их перечень приведен в табл. 2 в нотации Manchester Syntax.

Табл. 2. Определяемые классы KazFakeOnto.

Класс	Определение
FakeNews	NewsItem and hasVeracityLabel value Fake
RealNews	NewsItem and hasVeracityLabel value Real
KazakhFakeNews	FakeNews and hasLanguage value Kazakh
RussianFakeNews	FakeNews and hasLanguage value Russian
ConspiracyNews	NewsItem and hasFakeNewsType value ConspiracyTheory

EmotionallyManipulativeNews	NewsItem and usesTechnique value Emotional-Pressure
MachineGeneratedNews	NewsItem and generatedBy some Generator
EvidentialFake	FakeNews and containsMarker some EvidentialMarker
HighConfidenceFake	FakeNews and annotatorConfidence some xsd:integer[>= 4]
ReliableRealNews	RealNews and hasSource some (hasCredibility value HighCredibility)
MultimodalNews	NewsItem and hasModality min 2 Modality

Аннотатор заполняет только элементарные уровни разметки – язык, достоверность, тип, технику, модальность, источник, уверенность; все производные категории вычисляются автоматически. Так, EvidentialFake объединяет фейковые материалы с эвиденциальным маркером; ReliableRealNews требует источника с высоким уровнем доверия, т. е. предполагает обход графа на два шага; HighConfidenceFake использует фасет xsd:integer[>= 4]; MultimodalNews – квалифицированное ограничение мощности. Ни одна из этих категорий не хранится в разметке, и при изменении элементарных значений все производные пересчитываются согласованно.

Приведем фрагмент онтологии в синтаксисе Turtle, описывающий один материал корпуса.

```

kf:N001 a owl:NamedIndividual , kf:NewsItem ;
    kf:newsId "KFC-0001" ; kf:hasLanguage kf:Kazakh ; kf:hasVeracityLabel kf:Fake ;
    kf:hasFakeNewsType kf:ConspiracyTheory ; kf:usesTechnique kf:EmotionalPressure ;
    kf:hasAuthorIntent kf:IdeologicalIntent ; kf:hasModality kf:TextModality ;
    kf:hasSource kf:AnonymousTelegramChannel ; kf:containsMarker kf:Marker_MYS ;
    kf:annotatorConfidence 5 ; kf:tokenCount 148 ; kf:partOfCorpus kf:KazFakeCorpus .

```

Материал kf:N001 не отнесен явно ни к одному определяемому классу; ризонер выводит его принадлежность к FakeNews, KazakhFakeNews, ConspiracyNews, EmotionallyManipulativeNews, EvidentialFake и HighConfidenceFake.

Доступность и проверка непротиворечивости. Онтология распространяется вместе с сопроводительными материалами исследования в двух сериализациях – RDF/XML и Turtle – на условиях лицензии Creative Commons Attribution 4.0 International (CC BY 4.0).

Логический вывод выполнялся ризонером HermiT 1.4.3 [23]. Проверка непротиворечивости не выявила неудовлетворимых классов: ни один именованный класс не отнесен к owl:Nothing. Все одиннадцать определяемых классов получили непустые множества экземпляров. В табл. 3 представлены результаты для девяти классов, имеющих самостоятельное значение для демонстрации возможностей вывода; не включенные в таблицу классы RussianFakeNews и EmotionallyManipulativeNews содержат соответственно 6 и 4 экз. Отметим деталь реализации: даты представлены типом xsd:dateTime, поскольку xsd:date не входит в обязательную карту типов данных OWL 2 и ризонером HermiT в применявшейся конфигурации не обрабатывался.

5. ЗАПРОСНЫЙ СЛОЙ И АПРОБАЦИЯ

Работоспособность модели проверялась SPARQL-запросами [24] к утвержденным триплетам и DL-запросами к выведенной модели: вкладка SPARQL в Protégé работает без учета вывода, поэтому определяемые классы проверяются DL-запросами при запущенном ризонере. Все результаты получены на демонстрационном фрагменте из 18 материалов и характеризуют состав фрагмента, а не корпуса в целом.

Анализ разметки средствами SPARQL. Пятнадцать запросов Q1–Q15 покрывают четыре группы задач: описательную статистику корпуса (распределение по языку и достоверности, частотность техник, нагрузка аннотаторов, метаданные); анализ разметки (профиль материала, мультимодальные фейки, фейки из ненадежных источников, достоверные новости и их подтверждения); лингвистический слой (формы эвиденциальных маркеров и их распределение по классам достоверности); слой моделей детекции (сравнение моделей, машинно-по-

рожденные материалы и их исходные тексты, объяснения, расхождения разметки и предсказаний).

Особый интерес представляет запрос Q5 о распределении эвиденциальных маркеров по классам достоверности:

```
SELECT ?veracity (COUNT(DISTINCT ?n) AS ?withMarker) WHERE {
  ?v a kf:Veracity ; rdfs:label ?veracity . OPTIONAL { ?n kf:hasVeracityLabel ?v ;
    kf:containsMarker ?m } FILTER(lang(?veracity) = "ru") }
GROUP BY ?veracity
```

Конструкция OPTIONAL здесь существенна: без нее класс, ни один материал которого не содержит маркера, не попал бы в результат, и нулевое значение было бы неотличимо от отсутствия материалов такого класса. Запрос возвращает две строки – «фейковая: 7» и «достоверная: 0».

Запрос Q4 детализирует картину по формам маркеров: в казахском материале зафиксированы аффикс *-мыс/-міс*, частица *екен*, ренарративные конструкции *деп айтады/дейді* и *көрінеді*; в русском – *якобы*, *по слухам*, *как сообщается*. Различие грамматических средств выражения эвиденциальности делает этот уровень разметки типологически значимым и объясняет введение подклассов класса EvidentialMarker. Наблюдаемая асимметрия является следствием инструкций порождения; воспроизводится ли она на естественно возникающей дезинформации – открытый вопрос.

Логический вывод: DL-запросы. Результаты тринадцати DL-запросов приведены в табл. 3. Ни одна из проверяемых категорий не утверждается в онтологии явно.

Табл. 3. DL-запросы и результаты вывода.

№	Запрос (Manchester Syntax)	Количество экз.
D1	FakeNews	12
D2	RealNews	6
D3	KazakhFakeNews	6
D4	MachineGeneratedNews	3
D5	EvidentialFake	7

D6	HighConfidenceFake	9
D7	ReliableRealNews	5
D8	MultimodalNews	6
D9	FakeNews and (usesTechnique value Clickbait)	3
D10	NewsItem and (annotatedBy value Annotator1)	4
D11	FakeNews and (hasSource some (hasCredibility value LowCredibility))	12
D12	NewsItem and (containsMarker some EvidentialAffix)	1
D13	ConspiracyNews	2

Наиболее показательны три результата. Запрос D7 возвращает пять материалов из шести достоверных: один исключается, поскольку его источник имеет средний уровень доверия, вывод опирается на обход графа через сущность Source. Запрос D10 возвращает материалы, размеченные конкретным аннотатором, при том что свойство annotatedBy нигде не утверждено. Запрос D13 объединяет два материала конспирологического типа, один из которых порожден языковой моделью из второго.

Отдельного комментария требует результат D11: все двенадцать материалов класса FAKE во фрагменте связаны с источниками низкого доверия. Это следствие устройства фрагмента, а не наблюдаемая закономерность; в корпусе в целом уровень доверия к источнику не должен совпадать с меткой достоверности, иначе он превращается в признак, тривиально предсказывающий класс.

Онтология как слой над результатами моделей детекции. Онтология охватывает не только эталонную разметку, но и результаты систем автоматического обнаружения: классы DetectionModel, Prediction и Explanation связывают материал, предсказанную метку, уверенность модели и порожденные объяснения. Метрики моделей, разработанных и оцененных на KazFakeCorpus в рамках проекта AP26195591, представлены свойствами accuracy, macroF1, f1Russian, f1Kazakh и извлекаются запросом Q7 (табл. 4).

Табл. 4. Модели детекции, представленные в онтологии (значения в %).

Модель	Accuracy	Macro-F1	F1 (рус.)	F1 (каз.)
ADVISOR (RAG + двунаправленное кросс-внимание)	87.9	88.0	88.7	86.9
mDeBERTa-v3	84.5	84.5	–	–
multilingual-E5-large	83.3	83.3	–	–
XLM-RoBERTa [25]	82.4	82.4	–	–

Примечание. Приведенные значения получены в рамках проекта AP26195591 на корпусе KazFakeCorpus; описание архитектуры модели ADVISOR и протокола ее оценки готовится к публикации отдельно и здесь не является предметом самостоятельной экспериментальной оценки.

Различие качества по языкам для лучшей модели (86.9 против 88.7) составляет 1.8 п. п. и получено в одном прогоне без оценки доверительных интервалов, поэтому не может служить основанием для содержательных выводов. Как рабочую гипотезу отметим возможную связь с наблюдением по запросу Q4: значительная часть маркеров недостоверности в казахском языке выражается грамматически, что может затруднять их усвоение многоязычными энкодерами.

Запрос Q11 автоматизирует анализ ошибок: он отбирает материалы, для которых предсказанная метка расходится с эталонной, и возвращает обе метки, модель, ее уверенность и технику дезинформации. Во фрагменте единственная ошибка приходится на материал с техникой преуменьшения, ошибочно классифицированный моделью XLM-RoBERTa как достоверный с уверенностью 0.63.

Запрос Q10 извлекает три объяснения, порожденные объяснимой моделью для одного материала: на здравом смысле (E_CS), стилистико-текстовое (E_TX) и на внешних свидетельствах (E_DB). Для материала KFC-0001 при уверенности модели 0.94 объяснение E_TX прямо апеллирует к эвиденциальному аффиксу *-мыс/-міс*, т. е. к признаку, зафиксированному в онтологии на уровне разметки. Возможность сопоставить объяснение модели с независимо размеченным лингвистическим признаком служит практическим инструментом оценки достоверности объяснений.

6. ОБСУЖДЕНИЕ

Онтология как спецификация схемы разметки. Формализация переводит часть инструкции для аннотатора в проверяемую форму: ограничения, ранее существовавшие как соглашения – единственность метки достоверности, единственность уровня доверия источника, несовместимость классов REAL и FAKE, становятся аксиомами, нарушение которых обнаруживается автоматически. Для схем, развиваемых на протяжении нескольких итераций, это снижает риск скрытого рассогласования версий разметки. Требования полноты средствами OWL не проверяются и требуют ограничений SHACL.

Разделение наблюдаемого и выводимого. Принцип «аннотатор размечает только элементарные уровни, содержательные категории вычисляются» ограничивает разметку теми решениями, которые эксперт действительно принимает при чтении текста, и исключает класс ошибок, связанных с ручным дублированием производных меток. Одновременно он делает исследовательские категории пересматриваемыми: изменение определения категории требует правки одной аксиомы, а не переразметки корпуса. Существенно и то, что статистика корпуса, лингвистическая гипотеза об эвиденциальности, анализ ошибок классификатора и оценка объяснений формулируются над одним графом.

Ограничения. Основное ограничение унаследовано от корпуса: материалы класса FAKE синтетические и могут отличаться от естественно возникающей дезинформации стилистически и дискурсивно; контролируемая трансформация обеспечивает сопоставимость классов, но создает риск того, что обучаемая модель усваивает признаки порождающей модели. Свойства `generatedBy` и `derivedFromSeed` введены именно для того, чтобы этот риск можно было измерять. Второе ограничение модельное: уровни разметки заданы как свойства материала, а не объекта аннотации, поэтому двойная разметка и процедура арбитража в графе не представлены. Третье: в онтологию перенесен иллюстративный фрагмент, а не весь корпус. Четвертое связано с производительностью: вывод над графом из нескольких десятков тысяч индивидов требует перехода к триплетному хранилищу, причем профили OWL, поддерживаемые правилами машинами, не покрывают фасеты типов данных и квалифицированные ограничения мощности; практическим решением представляется предвычисление содержательных категорий запросами CONSTRUCT.

Этические аспекты и условия использования. Ресурс содержит машинно-порожденные тексты, воспроизводящие приемы дезинформации, что создает риск нецелевого использования. Материалы класса REAL – это находящиеся в открытом доступе официальные публикации; персональные данные частных лиц из текстов удалялись. Материалы класса FAKE являются синтетическими обучающими примерами и не предназначены для представления в качестве достоверных сообщений.

Доступ к текстовой части корпуса предоставляется для научных и образовательных целей по обоснованному запросу к авторам проекта и при условии принятия пользовательского соглашения, которое запрещает применение материалов для введения пользователей в заблуждение, имитации достоверных новостей и распространения дезинформации. Пользователь обязуется обеспечивать безопасное хранение данных, не публиковать синтетические тексты без указания на их искусственное происхождение и приводить ссылку на KazFakeCorpus.

Направления развития. Ближайшими задачами являются: перенос полного корпуса в онтологию, перенос уровней разметки на объект аннотации и автоматическая генерация интерфейса аннотатора из справочников онтологии. Содержательные направления – это сопоставление естественной и порожденной дезинформации в терминах определяемых классов и расширение типологии эвиденциальных средств казахского языка, а также согласование KazFakeOnto со словарем ClaimReview [14] и графами ClaimsKG [15] и CimpleKG [16].

ЗАКЛЮЧЕНИЕ

Представлен сбалансированный двуязычный корпус KazFakeCorpus объемом 4276 текстов, размеченный по многоуровневой семантической схеме, которая охватывает достоверность сообщения, тип фейкового контента, технику дезинформации, коммуникативное намерение, модальность и характеристики источника; согласованность разметки по коэффициенту Krippendorff's Alpha составила 0.79–0.88.

Основной вклад статьи состоит в переводе схемы разметки в формальную онтологическую модель KazFakeOnto на языке OWL 2 DL: 38 классов, 26 объектных и 22 свойства-значения данных, 112 индивидов и одиннадцать определяе-

мых классов, выводимых ризонером из элементарных уровней разметки. Непротиворечивость подтверждена ризонером *HermiT*, запросный слой апробирован на пятнадцати SPARQL- и тринадцати DL-запросах.

Предложенный подход моделирует фейковую новость как многомерное семантическое явление и обеспечивает воспроизводимую основу для исследований в области анализа дезинформации и кросс-языкового обнаружения фейковых новостей для малоресурсных языков.

Благодарности

Исследование выполнено при поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан (грант № AP26195591).

СПИСОК ЛИТЕРАТУРЫ

1. *Zhou X., Zafarani R.* A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities // *ACM Computing Surveys*. 2020. Vol. 53, No. 5. P. 1–40. <https://doi.org/10.1145/3395046>
2. *Amjad M., Sidorov G., Zhila A., Gómez-Adorno H., Voronkov I., Gelbukh A.* “Bend the Truth”: Benchmark Dataset for Fake News Detection in Urdu Language and Its Evaluation // *Journal of Intelligent & Fuzzy Systems*. 2020. Vol. 39, No. 2. P. 2457–2469. <https://doi.org/10.3233/JIFS-179905>
3. *Azad R., Mohammed B., Mahmud R., Zrar L., Sdiqa S.* Fake News Detection in Low-Resourced Languages “Kurdish Language” Using Machine Learning Algorithms // *Turkish Journal of Computer and Mathematics Education*. 2021. Vol. 12, No. 6. P. 4219–4225.
4. *Makhambetov O., Makazhanov A., Yessenbayev Z., Matkarimov B., Sabirgaliyev I., Sharafudinov A.* Assembling the Kazakh Language Corpus // *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*. 2013. P. 1022–1031. URL: <https://aclanthology.org/D13-1104/> (дата обращения 15.08.2026).
5. *Sambetbayeva M., Nekessova A., Yerimbetova A., Bayangali A., Kaldarova M., Telman D., Smailov N.* A Multi-Level Annotation Model for Fake News Detection: Implementing Kazakh-Russian Corpus via Label Studio // *Big Data and Cognitive Computing*. 2025. Vol. 9, No. 8. Article 215. <https://doi.org/10.3390/bdcc9080215>

6. *Lamasheva Zh., Nekessova A., Kantureyeva M., Sambetbayeva M., Kaldarova M., Nazymkhan A.* KazFakeCorpus: A Bilingual Corpus with Multi-Level Semantic Annotation for Fake News Detection // *Big Data and Cognitive Computing*. 2026. Vol. 10, No. 6. Article 183. <https://doi.org/10.3390/bdcc10060183>

7. *Nekessova A., Sidorova E., Lamasheva Zh., Sambetbayeva M., Yerimbetova A., Kaldarova M., Nazymkhan A.* KazFakeCorpus: A Bilingual Multi-Level Corpus for Semantic Detection of Fake News // *Компьютерная лингвистика и интеллектуальные технологии: по материалам международной конференции «Диалог»*. Вып. 24. М.: МАКС Пресс, 2026. С. 438–447. <https://doi.org/10.29003/2075-7182-2026-24-438-447>

8. *Shu K., Mahudeswaran D., Wang S., Lee D., Liu H.* FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media // *Big Data*. 2020. Vol. 8, No. 3. P. 171–188. <https://doi.org/10.1089/big.2020.0062>

9. *Li Y., He H., Bai J., Wen D.* MCFEND: A Multi-Source Benchmark Dataset for Chinese Fake News Detection // *Proc. of the ACM Web Conference 2024*. 2024. P. 4018–4027. <https://doi.org/10.1145/3589334.3645385>

10. *Ashraf S., Bezzaoui I., Andone I., Markowetz A., Fegert J., Flek L.* DeFaktS: A German Dataset for Fine-Grained Disinformation Detection through Social Media Framing // *Proceedings of LREC-COLING 2024*. 2024. P. 4580–4591.

URL: <https://aclanthology.org/2024.lrec-main.409/> (дата обращения 15.08.2026).

11. *Tandoc E.C., Lim Z.W., Ling R.* Defining “Fake News”: A Typology of Scholarly Definitions // *Digital Journalism*. 2018. Vol. 6, No. 2. P. 137–153. <https://doi.org/10.1080/21670811.2017.1360143>

12. *Da San Martino G., Barrón-Cedeño A., Wachsmuth H., Petrov R., Nakov P.* SemEval-2020 Task 11: Detection of Propaganda Techniques in News Articles // *Proceedings of the Fourteenth Workshop on Semantic Evaluation (SemEval-2020)*. 2020. P. 1377–1414. <https://doi.org/10.18653/v1/2020.semeval-1.186>

13. *Bezzaoui I., Fafalios P., Fegert J., Rettinger A., Todorov K.* Decoding Deception with TAXODIS – a Taxonomy of Disinformation Cues for Fine-Grained Text Labeling // *Semantic Web Journal*. URL: <https://zenodo.org/records/14264593> (дата обращения 15.08.2026).

14. ClaimReview // *Schema.org*. URL: <https://schema.org/ClaimReview> (дата

обращения 15.08.2026).

15. *Tchechmedjiev A., Fafalios P., Boland K., Gasquet M., Zloch M., Zapilko B., Dietze S., Todorov K.* ClaimsKG: A Knowledge Graph of Fact-Checked Claims // *The Semantic Web – ISWC 2019. Lecture Notes in Computer Science. Vol. 11779.* Cham: Springer, 2019. P. 309–324. https://doi.org/10.1007/978-3-030-30796-7_20

16. *Burel G., Mensio M., Peskine Y., Troncy R., Papotti P., Alani H.* CimpleKG: A Continuously Updated Knowledge Graph on Misinformation, Factors and Fact-Checks // *The Semantic Web – ISWC 2024. Lecture Notes in Computer Science. Vol. 15233.* Cham: Springer, 2025. P. 97–114. https://doi.org/10.1007/978-3-031-77847-6_6

17. *Bani-Hani A., Adedugbe O., Benkhelifa E., Majdalawieh M.* Fandet Semantic Model: An OWL Ontology for Context-Based Fake News Detection on Social Media // *Combating Fake News with Computational Intelligence Techniques. Studies in Computational Intelligence. Vol. 1001.* Cham: Springer, 2022. P. 91–125. https://doi.org/10.1007/978-3-030-90087-8_5

18. *Abdygalym B.Kh., Adali E., Sambetbayeva M.A., Sadirmekova Zh.B., Nazymkhan A.A.* A Conceptual Model for Ontology-Based Detection of Information Operations in Digital Media // *Bulletin of Shakarim University. Technical Sciences. 2025. No. 4(20).* P. 36–44. [https://doi.org/10.53360/2788-7995-2025-4\(20\)-5](https://doi.org/10.53360/2788-7995-2025-4(20)-5)

19. *Елизаров А.М., Кириллович А.В., Липачев Е.К., Невзорова О.А.* Цифровая экосистема OntoMath как подход к построению пространства математических знаний // *Электронные библиотеки. 2023. Т. 26, № 2.* С. 154–202. <https://doi.org/10.26907/1562-5419-2023-26-2-154-202>

20. *Krippendorff K.* Content Analysis: An Introduction to Its Methodology. 4th ed. Thousand Oaks: SAGE Publications, 2018.

21. OWL 2 Web Ontology Language Document Overview (Second Edition). W3C Recommendation, 11 December 2012. URL: <https://www.w3.org/TR/owl2-overview/> (дата обращения 15.08.2026).

22. *Aikhenvald A.Y.* Evidentiality. Oxford: Oxford University Press, 2004.

23. *Glimm B., Horrocks I., Motik B., Stoilos G., Wang Z.* HermiT: An OWL 2 Reasoner // *Journal of Automated Reasoning. 2014. Vol. 53, No. 3.* P. 245–269. <https://doi.org/10.1007/s10817-014-9305-1>

24. SPARQL 1.1 Query Language. W3C Recommendation, 21 March 2013. URL: <https://www.w3.org/TR/sparql11-query/> (дата обращения 15.08.2026).

25. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzmán F., Grave E., Ott M., Zettlemoyer L., Stoyanov V. Unsupervised Cross-Lingual Representation Learning at Scale // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. P. 8440–8451.

<https://doi.org/10.18653/v1/2020.acl-main.747>

FROM AN ANNOTATED CORPUS TO AN ONTOLOGY: A FORMAL MODEL OF MULTI-LEVEL SEMANTIC ANNOTATION OF FAKE NEWS IN THE KAZAKH–RUSSIAN MEDIA SPACE (KazFakeCorpus / KazFakeOnto)

A. A. Nekessova¹ [0009-0001-7642-0648], E. A. Sidorova² [0000-0001-8731-3058],
Zh. B. Lamasheva³ [0000-0002-9535-2636], M. A. Sambetbayeva⁴ [0000-0001-9358-1614],
A. S. Yerimbetova⁵ [0000-0002-2013-1513], M. Zh. Kaldarova⁶ [0000-0001-7494-9794],
A. A. Nazymkhan⁷ [0009-0007-4366-6709]

^{1, 3, 4, 6, 7}International Science Complex Astana, Astana, Kazakhstan

^{1, 3, 4, 7}L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

²A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

⁵Institute of Information and Computational Technologies CS MSHE RK, Almaty, Kazakhstan

⁵School of Engineering and Information Technology, META University, Almaty, Kazakhstan

⁶Astana International University, Astana, Kazakhstan

¹necessoaa_aa_1@enu.kz, ²lsidorova@iis.nsk.su, ³lamasheva_zhb@enu.kz,

⁴sambetbayeva_ma_1@enu.kz, ⁵aigerian8888@gmail.com,

⁶mira_kaldarova@aiu.edu.kz, ⁷nazymkhan_aa@enu.kz

Abstract

The paper addresses the formal representation of knowledge about disinformation in a low-resource bilingual media environment. Automatic fake news detection is traditionally formulated as a binary REAL/FAKE classification task, which does not capture the internal structure of an unreliable message: the type of fake content, the disinformation technique employed, the author's communicative intent, and the

characteristics of the source and evidence base.

Two interrelated resources are presented. The first is KazFakeCorpus, a bilingual corpus of 4276 texts in Kazakh and Russian, balanced by class within each language; the REAL class was drawn from official materials of the Gov.kz portal and the FAKE class from their controlled transformations. The texts were annotated in Label Studio by two independent experts following a multi-level semantic scheme; Krippendorff's Alpha for the principal levels ranged from 0.79 to 0.88. The second resource is KazFakeOnto, an ontology that recasts the annotation scheme from a flat relational representation into a formal OWL 2 DL model: 38 classes, 26 object properties, 22 data properties, 112 individuals and 1155 asserted triples. Eleven defined classes are not asserted manually but inferred by a reasoner from property values; a property chain links a news item to its annotator through an annotation object. Consistency was verified with HermiT, and the query layer was evaluated on fifteen SPARQL and thirteen DL queries.

The ontological layer is shown to provide a reproducible specification of the annotation scheme, the automatic inference of semantic categories, and a unified formal environment for formulating corpus-linguistic questions and analysing machine-learning model errors. The limitations of the approach and directions for scaling the ontology to the full corpus and to naturally occurring disinformation are discussed.

Keywords: *fake news detection, ontology, OWL 2, SPARQL, logical inference, bilingual corpus, Kazakh language, multi-level annotation, disinformation techniques, explainable NLP, low-resource languages, KazFakeCorpus, KazFakeOnto.*

Acknowledgements. This research was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant No. AP26195591).

REFERENCES

1. Zhou X., Zafarani R. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities // ACM Computing Surveys. 2020. Vol. 53, No. 5. P. 1–40. <https://doi.org/10.1145/3395046>
2. Amjad M., Sidorov G., Zhila A., Gómez-Adorno H., Voronkov I., Gelbukh A. “Bend the Truth”: Benchmark Dataset for Fake News Detection in Urdu Language and Its Evaluation // Journal of Intelligent & Fuzzy Systems. 2020. Vol. 39, No. 2. P. 2457–

2469. <https://doi.org/10.3233/JIFS-179905>

3. Azad R., Mohammed B., Mahmud R., Zrar L., Sdiqa S. Fake News Detection in Low-Resourced Languages “Kurdish Language” Using Machine Learning Algorithms // Turkish Journal of Computer and Mathematics Education. 2021. Vol. 12, No. 6. P. 4219–4225.

4. Makhambetov O., Makazhanov A., Yessenbayev Z., Matkarimov B., Sab-yrgaliyev I., Sharafudinov A. Assembling the Kazakh Language Corpus // Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2013). 2013. P. 1022–1031. URL: <https://aclanthology.org/D13-1104/> (дата обращения 15.08.2026).

5. Sambetbayeva M., Nekessova A., Yerimbetova A., Bayangali A., Kaldarova M., Telman D., Smailov N. A Multi-Level Annotation Model for Fake News Detection: Implementing Kazakh-Russian Corpus via Label Studio // Big Data and Cognitive Computing. 2025. Vol. 9, No. 8. Article 215. <https://doi.org/10.3390/bdcc9080215>

6. Lamasheva Zh., Nekessova A., Kantureyeva M., Sambetbayeva M., Kaldarova M., Nazymkhan A. KazFakeCorpus: A Bilingual Corpus with Multi-Level Semantic Annotation for Fake News Detection // Big Data and Cognitive Computing. 2026. Vol. 10, No. 6. Article 183. <https://doi.org/10.3390/bdcc10060183>

7. Nekessova A., Sidorova E., Lamasheva Zh., Sambetbayeva M., Yerimbetova A., Kaldarova M., Nazymkhan A. KazFakeCorpus: A Bilingual Multi-Level Corpus for Semantic Detection of Fake News // Компьютерная лингвистика и интеллектуальные технологии: по материалам международной конференции «Диалог». Вып. 24. М.: МАКС Пресс, 2026. С. 438–447. <https://doi.org/10.29003/2075-7182-2026-24-438-447>

8. Shu K., Mahudeswaran D., Wang S., Lee D., Liu H. FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media // Big Data. 2020. Vol. 8, No. 3. P. 171–188. <https://doi.org/10.1089/big.2020.0062>

9. Li Y., He H., Bai J., Wen D. MCFEND: A Multi-Source Benchmark Dataset for Chinese Fake News Detection // Proc. of the ACM Web Conference 2024. 2024. P. 4018–4027. <https://doi.org/10.1145/3589334.3645385>

10. Ashraf S., Bezzaoui I., Andone I., Markowetz A., Fegert J., Flek L. DeFaktS: A

German Dataset for Fine-Grained Disinformation Detection through Social Media Framing // Proceedings of LREC-COLING 2024. 2024. P. 4580–4591.

URL: <https://aclanthology.org/2024.lrec-main.409/> (дата обращения 15.08.2026).

11. *Tandoc E.C., Lim Z.W., Ling R.* Defining “Fake News”: A Typology of Scholarly Definitions // Digital Journalism. 2018. Vol. 6, No. 2. P. 137–153.

<https://doi.org/10.1080/21670811.2017.1360143>

12. *Da San Martino G., Barrón-Cedeño A., Wachsmuth H., Petrov R., Nakov P.* SemEval-2020 Task 11: Detection of Propaganda Techniques in News Articles // Proceedings of the Fourteenth Workshop on Semantic Evaluation (SemEval-2020). 2020. P. 1377–1414. <https://doi.org/10.18653/v1/2020.semeval-1.186>

13. *Bezzaoui I., Fafalios P., Fegert J., Rettinger A., Todorov K.* Decoding Deception with TAXODIS – a Taxonomy of Disinformation Cues for Fine-Grained Text Labeling // Semantic Web Journal. URL: <https://zenodo.org/records/14264593> (дата обращения 15.08.2026).

14. ClaimReview // Schema.org. URL: <https://schema.org/ClaimReview> (дата обращения 15.08.2026).

15. *Tchechmedjiev A., Fafalios P., Boland K., Gasquet M., Zloch M., Zapilko B., Dietze S., Todorov K.* ClaimsKG: A Knowledge Graph of Fact-Checked Claims // The Semantic Web – ISWC 2019. Lecture Notes in Computer Science. Vol. 11779. Cham: Springer, 2019. P. 309–324. https://doi.org/10.1007/978-3-030-30796-7_20

16. *Burel G., Mensio M., Peskine Y., Troncy R., Papotti P., Alani H.* CimpleKG: A Continuously Updated Knowledge Graph on Misinformation, Factors and Fact-Checks // The Semantic Web – ISWC 2024. Lecture Notes in Computer Science. Vol. 15233. Cham: Springer, 2025. P. 97–114. https://doi.org/10.1007/978-3-031-77847-6_6

17. *Bani-Hani A., Adedugbe O., Benkhelifa E., Majdalawieh M.* Fandet Semantic Model: An OWL Ontology for Context-Based Fake News Detection on Social Media // Combating Fake News with Computational Intelligence Techniques. Studies in Computational Intelligence. Vol. 1001. Cham: Springer, 2022. P. 91–125. https://doi.org/10.1007/978-3-030-90087-8_5

18. *Abdygalym B.Kh., Adali E., Sambetbayeva M.A., Sadirmekova Zh.B., Nazymkhan A.A.* A Conceptual Model for Ontology-Based Detection of Information Operations in Digital Media // Bulletin of Shakarim University. Technical Sciences. 2025. No. 4(20). P. 36–44. [https://doi.org/10.53360/2788-7995-2025-4\(20\)-5](https://doi.org/10.53360/2788-7995-2025-4(20)-5)

19. Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A. Tsifrovaya ekosistema OntoMath kak podkhod k postroeniyu prostranstva matematicheskikh znaniy [Digital Ecosystem OntoMath as an Approach to Building the Space of Mathematical Knowledge] // Russian Digital Libraries Journal. 2023. Vol. 26, No. 2. P. 154–202. <https://doi.org/10.26907/1562-5419-2023-26-2-154-202> (In Russ.)

20. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 4th ed. Thousand Oaks: SAGE Publications, 2018.

21. OWL 2 Web Ontology Language Document Overview (Second Edition). W3C Recommendation, 11 December 2012. URL: <https://www.w3.org/TR/owl2-overview/> (дата обращения 15.08.2026).

22. Aikhenvald A.Y. Evidentiality. Oxford: Oxford University Press, 2004.

23. Glimm B., Horrocks I., Motik B., Stoilos G., Wang Z. HermiT: An OWL 2 Reasoner // Journal of Automated Reasoning. 2014. Vol. 53, No. 3. P. 245–269. <https://doi.org/10.1007/s10817-014-9305-1>

24. SPARQL 1.1 Query Language. W3C Recommendation, 21 March 2013. URL: <https://www.w3.org/TR/sparql11-query/> (дата обращения 15.08.2026).

25. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzmán F., Grave E., Ott M., Zettlemoyer L., Stoyanov V. Unsupervised Cross-Lingual Representation Learning at Scale // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. P. 8440–8451.

СВЕДЕНИЯ ОБ АВТОРАХ



НЕКЕСОВА Анаргуль Аймуратовна – докторант Евразийского национального университета имени Л.Н. Гумилева и младший научный сотрудник Международного научного комплекса «Астана». Область научных интересов: обработка естественного языка, обнаружение дезинформации, корпусная лингвистика для малоресурсных языков.

Anargul A. NEKESOVA – PhD student, L.N. Gumilyov Eurasian National University and junior researcher at International Science Complex Astana. Research interests: natural language processing, disinformation detection, corpus linguistics for low-resource languages.

email: nekessova_aa_1@enu.kz

ORCID: 0009-0001-7642-0648

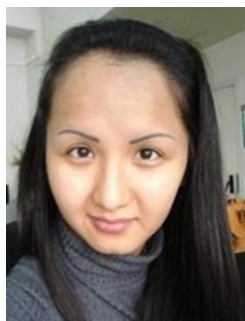


СИДОРОВА Елена Анатольевна – кандидат физико-математических наук, старший научный сотрудник Института систем информатики имени А.П. Ершова СО РАН. Область научных интересов: обработка естественного языка, онтологии и представление знаний, автоматический анализ текстов.

Elena A. SIDOROVA – PhD, senior researcher at the A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences. Research interests: natural language processing, ontologies and knowledge representation, automatic text analysis.

email: lsidorova@iis.nsk.su

ORCID: 0000-0001-8731-3058



ЛАМАШЕВА Жанар Бейбутовна – PhD, ассоциированный профессор Евразийского национального университета имени Л. Н. Гумилева и старший научный сотрудник Международного научного комплекса «Астана». Область научных интересов: компьютерная лингвистика, анализ медиатекстов, многоязычные языковые ресурсы.

Zhanar B. LAMASHEVA – PhD, associate professor at L. N. Gumilyov Eurasian National University and senior researcher at International Science Complex Astana. Research interests: computational linguistics, media text analysis, multilingual language resources.

email: lamasheva_zhb@enu.kz

ORCID: 0000-0002-9535-2636



САМБЕТБАЕВА Мадина Аралбаевна – PhD, ассоциированный профессор Евразийского национального университета имени Л. Н. Гумилева, ведущий научный сотрудник Международного научного комплекса «Астана». Область научных интересов: информационные системы, графы знаний, многоязычные системы обработки естественного языка.

Madina A. SAMBETBAYEVA – PhD, associate professor at L. N. Gumilyov Eurasian National University, leading researcher at International Science Complex Astana. Research interests: information systems, knowledge graphs, multilingual natural language processing systems.

email: sambetbayeva_ma_1@enu.kz

ORCID: 0000-0001-9358-1614



ЕРИМБЕТОВА Айгерим Сембековна – PhD, ведущий научный сотрудник Института информационных и вычислительных технологий КН МНВО РК, ассоциированный профессор Школы инженерии и информационных технологий Университета META. Область научных интересов: машинное обучение, обработка естественного языка, интеллектуальные информационные системы.

Aigerim S. YERIMBETOVA – PhD, leading researcher at the Institute of Information and Computational Technologies, associate professor at the School of Engineering and Information Technology, META University. Research interests: machine learning, natural language processing, intelligent information systems.

email: aigerian8888@gmail.com

ORCID: 0000-0002-2013-1513



КАЛДАРОВА Мира Жорабековна – PhD Международного университета Астана и старший научный сотрудник Международного научного комплекса «Астана». Область научных интересов: анализ данных, интеллектуальные системы обработки текстовой информации.

Mira Zh. KALDAROVA – PhD at Astana International University and senior researcher at International Science Complex Astana. Research interests: data analysis, intelligent text processing systems.

email: mira_kaldarova@aiu.edu.kz

ORCID: 0000-0001-7494-9794



НАЗЫМХАН Ақсауле Абзалқызы – инженер-программист Международного научного комплекса «Астана». Область научных интересов: обнаружение фейковых новостей, машинное обучение, аннотирование языковых данных.

Aksaule A. NAZYMKHAN – software engineer at International Science Complex Astana. Research interests: fake news detection, machine learning, language data annotation.

email: nazymkhan_aa@enu.kz

ORCID: 0009-0007-4366-6709

Материал поступил в редакцию 22 августа 2026 года