

АВТОМАТИЧЕСКОЕ ИЗВЛЕЧЕНИЕ АРГУМЕНТАТИВНЫХ ОТНОШЕНИЙ ИЗ ТЕКСТОВ НАУЧНОЙ КОММУНИКАЦИИ

Ю. А. Загорулько¹ [0000-0002-7111-6524], Е. А. Сидорова² [0000-0001-8731-3058],

И. Р. Ахмадеева³ [0000-0002-7371-1087]

^{1–3}Институт систем информатики им. А.П. Ершова СО РАН, г. Новосибирск,
Россия

¹zagor@iis.nsk.su, ²lsidorova @iis.nsk.su , ³i.r.akhmadeeva@iis.nsk.su

Аннотация

Сложность задачи извлечения аргументативных структур связана с такими проблемами, как выделение аргументативных сегментов, прогнозирование дальней связей между неконтактными сегментами, обучение на данных, размеченных с низкой степенью согласованности между аннотаторами. В настоящей работе рассмотрен подход к извлечению аргументативных отношений из достаточно больших текстов, относящихся к области научной коммуникации. Проведен сравнительный анализ методов тонкой настройки с использованием предобученной языковой модели типа Longformer, позволяющей учитывать длинные контексты, и двух методов, позволяющих учитывать расхождения аннотаторов в разметке аргументов за счет использования так называемых мягких меток, полученных путем равномерного сглаживания меток и усреднения экспертных оценок. Эксперименты проводились на четырех наборах данных, содержащих положительные и отрицательные примеры пар утверждений (посылка, заключение) и различающихся способами сегментации и средним размером текста. Наилучшие результаты получены на модели с усреднением экспертных оценок. В то же время отмечено, что модель, использующая сглаженные метки, также повышает точность классификаторов, но ухудшает полноту.

Ключевые слова: анализ аргументации, извлечение аргументативных отношений, научная коммуникация, проблемы сегментации, мягкая метка, сглаживание меток, языковая модель.

ВВЕДЕНИЕ

Аргументация является одной из конституирующих составляющих научной коммуникации (коммуникации на научные темы), поскольку содержанием научных и научно-популярных текстов является научно обоснованное знание. В научной сфере общения автор должен убедить в правомерности своих идей коллег-ученых, а в научно-популярной – широкую аудиторию.

При анализе аргументации, представленной в тексте, требуется не только извлекать аргументы и цепочки аргументов, подтверждающие или опровергающие некий тезис (абстрактная аргументация), но и исследовать структуру каждого аргумента, ее роль и значимость для всей аргументации в целом (структурная аргументация). Можно выделить два подхода к решению задачи автоматического извлечения структурированной аргументации:

— первый предполагает выполнение последовательности следующих шагов: сегментация текста → аргумент/не аргумент → простая структура аргумента → уточненная структура аргумента;

— второй подход, получивший название «все в одном» (all-in-one) включает всего два шага: сегментация текста → уточненная структура аргумента; он сразу решает задачу связанности выделенных сегментов аргументативными отношениями, и уже на этом основании решается вопрос об аргументативности утверждений и их типе.

Целью настоящей работы было экспериментальное исследование подхода all-in-one на основе нейросетевых моделей. Эксперименты проводились на двух русскоязычных корпусах текстов, относящихся к области научной коммуникации.

1. ОБЗОР РАБОТ

Для задачи интеллектуального анализа аргументации, как и для многих задач анализа естественного языка, характерен высокий уровень неоднозначности в размеченных обучающих данных. От такой неоднозначности сложно избавиться, поскольку она является неотъемлемым свойством самого языка. Поэтому в последнее время появились работы (см., например [1]), указывающие на необходимость учитывать неоднозначность при решении задач понимания естественного языка.

Одним из способов учета такой неоднозначности является сглаживание меток (label smoothing) – техника, широко используемая в глубоком обучении, которая, как было показано, улучшает качество работы модели при обучении на шумных данных [2] и предотвращает излишнюю самоуверенность (overconfidence) модели в ответе [3]. Эта техника вместо стандартного обучения с использованием прямого кодирования (one-hot encoding) предполагает использование сглаженных меток, полученных путем подмешивания равномерного вектора в исходный вектор меток.

Впервые техника сглаживания меток была предложена для решения задачи классификации изображений [4], но впоследствии была продемонстрирована ее применимость для задач анализа текста. Так, в работе [5] сглаживание меток применялось для калибровки модели при решении задачи автоматического определения логической связи между текстами (NLI), для которой характерна большая неоднозначность в разметке. Другой способ сглаживания меток заключается в усреднении меток, полученных от разных аннотаторов.

В настоящей работе исследована возможность применения техники сглаживания меток в новой области – интеллектуальном анализе аргументации.

2. КОРПУС ТЕКСТОВ С АРГУМЕНТАТИВНОЙ РАЗМЕТКОЙ

Используем два русскоязычных аннотированных корпуса текстов, представленных на платформе ArgNetBank Studio (<https://uniserv.iis.nsk.su/arg>). Научно-популярный корпус включает тексты двух подстилей, характеризующихся как тематической неоднородностью, так и жанровым разнообразием: научные новости и статьи с сайта habr.com/ru (habr-статьи). Корпус научной коммуникации [6] включает краткие научные статьи, научные обзоры и полноформатные научные статьи с комментариями рецензентов. В целом процесс разметки существенно зависит от жанра.

Короткие научные статьи, как правило, имеют один главный тезис, который достаточно легко выявить по позиции (два последних абзаца текста), и не имеют «длинных» связей. Разметка таких статей облегчается благодаря наличию в них элементов содержательной структуры (актуальность, цель работы, новизна) и формально-логической жанровой структуры (введение, заключение, обзор, методы, эксперимент, выводы).

В **рецензиях** главный тезис также находится в конце и сводится к трем вариантам: принять рецензируемую статью к публикации, принять после доработки или отклонить. Аргументативные отношения образуют кустовую структуру, аргументами первого порядка являются достоинства и/или недостатки статьи.

Главный тезис **научных новостей**, напротив, располагается в начале (лиде) новостной заметки; для заметки характерны повторы разной степени общности, имеется четкая жанровая структура с выделением хэдлиа, лида и бэкграунда.

В **статьях с комментариями рецензентов**, с одной стороны, как в обычной научной статье, можно опираться на цель, выделив главный тезис (один или несколько), с другой – комментарии образуют компонентный диалог, представляющий собой обмен стимулирующими и реагирующими репликами.

Набр-статьи могут вообще не иметь главного тезиса. Здесь можно опираться на содержательные разделы, а комментарии к текстам привносят в изложение черты каскадного диалога и полилога.

На основе анализа жанровых особенностей текстов было выделено два подкорпуса:

S-корпус – подкорпус коротких научных статей (в среднем 1053 словоупотребления);

L-корпус – подкорпус длинных (в среднем 3500 слов) текстов научной коммуникации, в котором собраны полноразмерные научные статьи, рецензии, аналитические научно-популярные статьи и новости о событиях в науке.

3. ПОДГОТОВКА НАБОРА ДАННЫХ

Для применения методов машинного обучения к задаче извлечения аргументативных связей необходимо:

а) создать на базе аннотированных текстов наборы данных, содержащие положительные и отрицательные примеры аргументативных отношений;

б) разработать механизм предварительного построения гипотез по заданному тексту, т. е. определить, каким образом текст будет разбиваться на утверждения и какие пары утверждений будут проверяться нашей моделью.

3.1. ПРОБЛЕМА СЕГМЕНТАЦИИ

Аргументативный анализ начинается с сегментации, т. е. разделения текста на осмысленные с точки зрения аргументации фрагменты (argumentative discourse units, ADUs). В общем случае это утверждения, в основе которых лежат пропозиции. Однако анализ сегментации, выполненной аннотаторами-людьми, показывает, что сегментами могут быть как более мелкие, так и более крупные фрагменты текста. При этом между аннотаторами далеко не всегда наблюдается согласие в отношении выполнения сегментации. В связи с этим трудно говорить о «золотом стандарте» в данном вопросе.

(S1) ChatGPT стал вторым чат-ботом, прошедшим широко известный Тест Тьюринга. (S2) Это значит, что во взаимодействии с ним судейской коллегии было невозможно определить общаются ли они с человеком или программой. (S3.1) Вдохновленные таким несомненным успехом, (S3.2) а также свободным доступом к боту предоставленном в OpenAI, (S3.3) многочисленные «уверовавшие в ИИ» начали наперебой предлагать приткнуть бота во все возможные ниши: от программирования до медицинских диагнозов. (S4) Даже поисковые системы забили тревогу в ожидании того, что бот подвинет их в предложении услуг поиска информации. (S5) На самом деле все эти ожидания не имеют под собой абсолютно никаких оснований. (S6) Ниже проиллюстрирую этот факт на конкретных примерах.

Рис. 1. Фрагмент текста из habr-статьи

В табл. 1 представлено распределение единиц рассуждения (ADUs), соответствующее аргументативно насыщенному абзацу из habr-статьи (см. рис. 1), который включили в анализ все аннотаторы (в эксперименте участвовали четыре аннотатора, обозначенных как A-1, A-2, A-3, A-4).

Табл. 1. Различия в сегментации фрагмента текста 4 аннотаторами

A-1	S1	S2	S3.1	S3.2	S3.3	S4		S5	
A-2	S1		S3.1	S3.2	S3.3	S4.1	S4.2	S5-S6	
A-3	S1	S2	S3					S5	S6
A-4	S1	S2	S3			S4		S5	

При этом A-2 исключил(а) из анализа второе предложение (S2), посчитав, что оно является не аргументирующим, а объяснительным, и предназначенным

для тех читателей, которые не знают, что такое тест Тьюринга. Но большинство аннотаторов посчитали S2 важным, обосновывающим S1 в качестве причины (A-4) или знака (A-1 и A-3). A-1 и A-4 удалили из анализа S6, возможно, сочли его метаязыковым, т. е. аргументативно незначимым. Последнее подтверждается включением S6 в комплекс с S5 аннотатором A-2, который посчитал(а) S6 индикатором примеров, представленных в тексте далее. Первые два аннотатора разбили S3, выделив самостоятельные сегменты S3.1 и S3.2 как основания для сегмента S3.3. S4 исключен из анализа аннотатором A-3, а A-2 разбил его на две клаузы, которые использованы в качестве поддерживающих посылок к S3.3.

Данный пример показывает, что, осуществляя аргументативную разметку, человек опирается на формальное выделение сегментов, но не ограничивается им, т. к. этот процесс скорее идет в ногу с процессом рассуждений относительно наличия и типа аргументативной связи. Усредненная оценка согласия между аннотаторами на данном фрагменте текста составляет 57.3%.

Оценка согласия сегментаций вычислялась в соответствии с алгоритмом, предложенным в работе [7]. Этот алгоритм основан на мере сходства множеств и дополнительно учитывает случаи, когда один сегмент пересекается с несколькими другими. Среднее значение согласия сегментаций на всем корпусе равно 61.2%, а коэффициенты каппа Козна (Cohen's k) и альфа Криппендорфа (Krippendorff's α) равны 0.42 и 0.43 соответственно.

Для автоматической сегментации текста помимо базового подхода — сегментации на предложения — использовались еще два подхода для более тонкой сегментации:

Q1: выделение клауз на основе синтаксического дерева предложения. Выделялись глагольные группы, причастные и деепричастные обороты;

Q2: выделение дискурсивных единиц на основе риторического анализа текста.

Сравнение работы сегментаторов с ручной сегментацией выявило существенные различия. Так, процент совпадения сегментов (на фрагментах с аргументацией) составил 49.5% и 46.75% для Q1 и Q2 соответственно. Кроме того, Q1 показал лучшие результаты на S-корпусе (56.2% против 44.9%), а Q2 — на L-корпусе (47% против 46.2%).

Анализ работы сегментаторов выявил следующие типичные ошибки: некорректная обработка разделителей предложений (точки или их отсутствие), неточное обращение с прямой речью (например, прямая речь со вставленной речью говорящего или косвенная речь с предшествующей речью говорящего), ошибочное разделение на клаузы при наличии свернутого предложения, представленного субстантивным предикатом, неверная идентификация разрывных сегментов и т. д.

3.2. ПОСТРОЕНИЕ ПАР УТВЕРЖДЕНИЙ ДЛЯ КЛАССИФИКАЦИИ

Для моделирования процесса построения гипотез все множество аргументативных связей было разделено на ближние (связь между двумя контактными сегментами) и дальние, соотношение которых в корпусе составило примерно 2 к 3. Для предсказания ближних связей рассматривались пары утверждений, получаемые с помощью скользящего окна. В качестве окна брались два подряд идущих сегмента (предложение и/или клауза). Для построения гипотез дальних связей механизм окна не подходил, поэтому были исследованы различные варианты взаиморасположения утверждений, относящихся к одному аргументу (посылка и заключение), друг относительно друга. В итоге для дальних связей отбирались утверждения, расположенные в границах одного предложения и располагающиеся либо в одном, либо в соседних абзацах (гипотеза компактности), что составило 52.9% от общего количества размеченных аннотаторами аргументативных связей.

Для обучения классификаторов было построено четыре набора данных, которые содержат положительные и отрицательные примеры пар утверждений (посылка, заключение). Пара считается положительным примером, если существует аргументативная связь между посылкой и заключением хотя бы в одной экспертной разметке. Отрицательные примеры отбирались следующим образом: для каждой позитивной пары утверждений генерировалась пара, в которой утверждения находятся в тех же абзацах (или в соседних, при отсутствии подходящих утверждений) и между которыми отсутствует путь от посылки к заключению во всех графах аргументации, соответствующих этому тексту.

На основе S-корпуса был создан один датасет, в котором утверждениями выступают предложения. На основе L-корпуса были созданы три датасета: в первом утверждениями также являются предложения, а в двух других — клаузы, полученные двумя различными методами сегментации (Q1 и Q2).

Каждому примеру были сопоставлены два типа меток: жесткие (one-hot encoding) и мягкие (soft labels). Применялись два способа получения мягких меток: а) равномерное сглаживание (label-smoothing), и б) усреднение (label-averaging). В случае равномерного сглаживания бинарные «жесткие» метки 0 и 1 заменялись на значения p и $1-p$ соответственно, где настраиваемый гиперпараметр p (коэффициент сглаживания меток) был подобран равным 0.1. При усреднении меток вес каждого примера определялся пропорционально доле аннотаторов, посчитавших данную пару утверждений связанной аргументативным отношением. Если в процессе усреднения для определенного текста существовала только одна разметка, то к полученным меткам применялся первый способ — равномерное сглаживание.

При построении тестовых датасетов в качестве положительных примеров отбирались только такие пары утверждений, которые посчитали аргументативными два и более аннотаторов. Пары утверждений, которые один из аннотаторов посчитал аргументативно связанными, а другой — нет, считались неоднозначными и в построении тестовых датасетов не участвовали.

4. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

В эксперименте использовался подход к решению задачи извлечения аргументативных отношений на основе тонкой настройки языковой модели (Fine Tuning). Задача извлечения аргументации all-in-one была сведена к предсказанию наличия или отсутствия связи между двумя утверждениями (предложениями, их частями или группами предложений). Другими словами, модель должна была решать задачу бинарной классификации: положительный ответ — связь есть, отрицательный ответ — связь отсутствует.

Предыдущие эксперименты показали, что в задаче предсказания аргументативной связи между утверждениями существенную роль играет контекст, в которых эти утверждения встречаются [8]. Поэтому были рассмотрены модели типа

Longformer [9], способные обрабатывать длинные фрагменты текста. По результатам предварительных экспериментов модель *kazzand/ru-longformer-large-4096* была выбрана для построения векторных представлений пар утверждений с учетом контекста их появления. Векторные представления утверждений строились путем усреднения векторов соответствующих токенов, а в качестве векторного представления контекста использовался вектор специального токена “[CLS]”. Конкатенация этих трех векторов подавалась в полносвязный слой для классификации на два класса: наличие или отсутствие аргументативного отношения.

В табл. 2 представлены результаты, полученные тремя моделями:

- 1) base model: базовая модель, обученная на жестких метках;
- 2) model-LS: модель, обученная на мягких метках, полученных из жестких меток с использованием сглаживания меток;
- 3) model-AA: модель, обученная на мягких метках, полученных путем усреднения экспертных оценок.

Все модели были обучены в течение 15 эпох со скоростью обучения $1e^{-5}$, коэффициент сглаживания меток для model-LS составил 0.1.

Табл. 2. Результаты предсказания наличия аргументативной связи

	S-корпус (предложения)			L-корпус (предложения)			L-корпус (клаузы-Q1)			L-корпус (клаузы-Q2)		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
base model	72.9	69.6	71.2	67.5	81	73.6	74.7	72.1	73.3	67.8	67.8	67.8
model-LS	73.5	80	76.6	72.6	60.5	66	75.9	65.9	70.6	68.4	57.8	62.7
model-AA	77.2	77.8	77.5	72.1	65.2	68.5	72.9	66.8	69.7	67.3	69.5	68.4

Как можно видеть, использование мягких меток повышает точность классификаторов, но для L-корпуса показатель полноты снижается. Это может быть связано с тем, что L-корпус характеризуется большим количеством текстов, размеченных только одним аннотатором. Для S-корпуса наилучшие результаты дала model-AA.

Более высокие показатели получены на S-корпусе, что может объясняться его гомогенностью, небольшой длиной текстов, относительной простотой аргументации и использованием аннотаторами единой методики разметки, что повысило уровень согласия между ними. Так, согласие аннотаторов на S-корпусе составило 64.2% для утверждений и 44.2% для связей, тогда как на L-корпусе – 58.9% и 33.2% соответственно.

ЗАКЛЮЧЕНИЕ

Мы рассмотрели подход к извлечению аргументативных отношений из русскоязычных текстов, обладающих достаточно сложной (для автоматической обработки) организационной структурой, свободным стилем изложения аргументов и достаточно большим объемом. Модель с мягкими метками повышает точность классификаторов, но ухудшает полноту. Наилучшие результаты были получены на модели с усреднением экспертных оценок.

В целом подход на основе мягких меток не показал значительных улучшений качества извлечения аргументации, поэтому дальнейшие работы мы планируем вести в направлении улучшения качества датасетов путем повышения согласия аннотаторов за счет строгого следования методическим рекомендациям и разработки правил унификации разметки.

Благодарности

Исследование выполнено при финансовой поддержке Российского научного фонда № 23-11-00261, <https://rscf.ru/project/23-11-00261/>.

СПИСОК ЛИТЕРАТУРЫ

1. *Meissner J.M., Thumwanit N., Sugawara S., Aizawa A.* Embracing Ambiguity: Shifting the Training Target of NLI Models // Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Online, August 2021. Association for Computational Linguistics: Vol. 2: Short Papers, P. 862–869.
<https://doi.org/10.18653/v1/2021.acl-short.109>

2. *Lukasik M., Bhojanapalli S., Menon A., Kumar S.* Does label smoothing mitigate label noise? // *Proceedings of the 37th International Conference on Machine Learning, Virtual*, 13–18 July 2020, Vol. 119, P. 6448–6458.

URL: <https://proceedings.mlr.press/v119/lukasik20a.html>

3. *Haque S., Bansal A., McMillan C.* Label smoothing improves neural source code summarization // *2023 IEEE/ACM 31st International Conference on Program Comprehension (ICPC)*, Melbourne, Australia, 15–16 May 2023. Institute of Electrical and Electronics Engineers: 2023, P. 101–112.

<https://doi.org/10.1109/ICPC58990.2023.00025>

4. *Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z.* Rethinking the Inception Architecture for Computer Vision // *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27–30 June, 2016. Institute of Electrical and Electronics Engineers: P. 2818–2826. <https://doi.org/10.1109/CVPR.2016.308>

5. *Wang Y., Wang M., Chen Y., Tao S., Guo J., Su C., Zhang M., Yang H.* Capture Human Disagreement Distributions by Calibrated Networks for Natural Language Inference // *Findings of the Association for Computational Linguistics: ACL 2022*, Dublin, Ireland. May 2022. Association for Computational Linguistics: 2022, P. 1524–1535. <https://doi.org/10.18653/v1/2022.findings-acl.120>

6. *Тимофеева М.К., Ильина Д.В., Кононенко И.С.* Аргументативная разметка корпуса текстов научной интернет-коммуникации: жанровый анализ и исследование типовых моделей рассуждения с помощью платформы ArgNetBank Studio // *Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация*. 2024. Т. 22, №1. С. 27–49. <https://doi.org/10.25205/1818-7935-2024-22-1-27-49>

7. *Shestakov V.K., Kononenko I.S., Sidorova E.A., Zagorulko Yu.A.* Assessing Inter-Annotator Agreement on Argumentative Markup // *2024 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. IEEE: 2024, P. 309–313. <https://doi.org/10.1109/SIBIRCON63777.2024.10758535>

8. *Akhmadeeva I., Sidorova E., Ilina D.* Argument mining in scientific communication: Comparative study // *Internet and modern society. Human-computer communication*. Cham: Springer Nature Switzerland, 2026. P. 152–166. https://doi.org/10.1007/978-3-031-96177-9_13

9. Beltagy I., Peters M. E., Cohan A. Longformer: The long-document transformer //arXiv preprint arXiv:2004.05150. 2020.

AUTOMATIC EXTRACTION OF ARGUMENTATIVE RELATIONS FROM SCIENTIFIC COMMUNICATION TEXTS

Yu. A. Zagorulko¹ [0000-0002-7111-6524], E. A. Sidorova² [0000-0001-8731-3058],

I. R. Akhmadeeva³ [0000-0002-7371-1087]

¹⁻³A.P. Ershov Institute of Informatics Systems of Siberian Branch of RAS, Novosibirsk, Russia

¹zagor@iis.nsk.su, ²lsidorova @iis.nsk.su , ³i.r.akhmadeeva@iis.nsk.su

Abstract

The complexity of the problem of extracting argumentative structures is associated with such problems as selecting argumentative segments, predicting long-range connections between non-contact segments, and training on data labeled with a low degree of inter-annotator consistency. In this paper, we consider an approach to extracting argumentative relations from fairly large texts related to scientific communication. A comparative analysis was performed of fine-tuning methods using a pre-trained Longformer-type language model that takes into account long contexts and two methods that take into account annotator discrepancies in argument labeling by using the so-called soft labels obtained by uniformly smoothing labels and averaging expert assessments. The experiments were conducted on four datasets containing positive and negative examples of statement pairs (premise, conclusion) and differing in segmentation methods and average text size. The best results were obtained using the model with averaging expert assessments. At the same time, it is noted that the model using smoothed labels also increases the accuracy of classifiers, but worsens the recall.

Keywords: *argument mining, argumentative relation extraction, scientific communication, segmentation problem, soft label, label smoothing, language model.*

REFERENCES

1. Meissner J.M., Thumwanit N., Sugawara S., Aizawa A. Embracing Ambiguity: Shifting the Training Target of NLI Models // Proceedings of the 59th Annual Meeting

of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Online, August 2021. Association for Computational Linguistics: Vol. 2: Short Papers, P. 862–869.

<https://doi.org/10.18653/v1/2021.acl-short.109>

2. *Lukasik M., Bhojanapalli S., Menon A., Kumar S.* Does label smoothing mitigate label noise? // *Proceedings of the 37th International Conference on Machine Learning, Virtual, 13–18 July 2020, Vol. 119, P. 6448–6458.*

URL: <https://proceedings.mlr.press/v119/lukasik20a.html>

3. *Haque S., Bansal A., McMillan C.* Label smoothing improves neural source code summarization // *2023 IEEE/ACM 31st International Conference on Program Comprehension (ICPC), Melbourne, Australia, 15–16 May 2023. Institute of Electrical and Electronics Engineers: 2023. P. 101–112.*

<https://doi.org/10.1109/ICPC58990.2023.00025>

4. *Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z.* Rethinking the Inception Architecture for Computer Vision // *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27–30 June, 2016. Institute of Electrical and Electronics Engineers: P. 2818–2826. https://doi.org/10.1109/CVPR.2016.308*

5. *Wang Y., Wang M., Chen Y., Tao S., Guo J., Su C., Zhang M., Yang H.* Capture Human Disagreement Distributions by Calibrated Networks for Natural Language Inference // *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland. May 2022. Association for Computational Linguistics: 2022, P. 1524–1535. https://doi.org/10.18653/v1/2022.findings-acl.120*

6. *Timofeeva M.K., Ilina D.V., Kononenko I.S.* Argumentative Annotation of the Scientific Internet-Communication Corpus: Genre Analysis and Study of Typical Reasoning Models based on the ArgNetBank Studio Platform // *NSU Vestnik. Series: Linguistics and Intercultural Communication. 2024. Vol. 22, No. 1. P. 27–49. (In Russ.) https://doi.org/10.25205/1818-7935-2024-22-1-27-49*

7. *Shestakov V.K., Kononenko I.S., Sidorova E.A., Zagorulko Yu.A.* Assessing Inter-Annotator Agreement on Argumentative Markup // *2024 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). IEEE: 2024, P. 309–313. https://doi.org/10.1109/SIBIRCON63777.2024.10758535*

8. *Akhmadeeva I., Sidorova E., Ilina D.* Argument mining in scientific communication: Comparative study // Internet and modern society. Human-computer communication. Cham: Springer Nature Switzerland, 2026. P. 152–166.

https://doi.org/10.1007/978-3-031-96177-9_13

9. *Beltagy I., Peters M. E., Cohan A.* Longformer: The long-document transformer //arXiv preprint arXiv:2004.05150. 2020.

СВЕДЕНИЯ ОБ АВТОРАХ

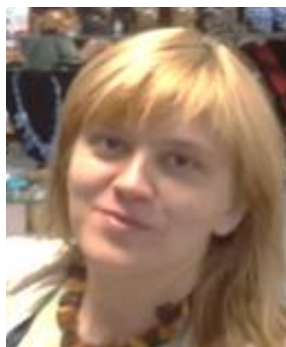


ЗАГОРУЛЬКО Юрий Алексеевич – кандидат технических наук, заведующий лабораторией Института систем информатики им. А.П. Ершова СО РАН, доцент кафедры программирования и кафедры систем информатики Новосибирского государственного университета. В списке научных трудов более 290 публикаций в области искусственного интеллекта, разработки интеллектуальных систем, инженерии знаний, онтологического моделирования и компьютерной лингвистики.

Yury Alekseevich ZAGORULKO – Cand. Sc. (Technology), Head of Laboratory at the A. P. Ershov Institute of Informatics Systems of Siberian Branch of RAS, an associate professor at Novosibirsk State University. List of scientific works includes more than 290 publications in the fields of AI, Knowledge and Ontology Engineering, Intelligent System Development and Computational Linguistics.

email: zagor@iis.nsk.su

ORCID: 0000-0002-7111-6524



СИДОРОВА Елена Анатольевна – кандидат физико-математических наук, старший научный сотрудник Института систем информатики им. А.П. Ершова СО РАН, доцент кафедры программирования и кафедры систем информатики Новосибирского государственного университета. В списке научных трудов более 160 работ в области компьютерной лингвистики, онтологического инжиниринга и разработки интеллектуальных систем.

Elena Anatolievna SIDOROVA – PhD (2006). She is a senior researcher at the A. P. Ershov Institute of Informatics Systems of Siberian Branch of RAS, an associate professor at Novosibirsk State University. List of scientific works includes more than 160 peer-reviewed publications in the fields of Computational Linguistics, Intelligent System Development, Knowledge and Ontology Engineering.

email: lsidorova@iis.nsk.su

ORCID: 0000-0001-8731-3058



АХМАДЕЕВА Ирина Равильевна – младший научный сотрудник Института систем информатики им. А.П. Ершова СО РАН, ассистент кафедры программирования Новосибирского государственного университета. В списке научных трудов около 30 работ в области искусственного интеллекта, разработки интеллектуальных систем и компьютерной лингвистики.

Irina Ravilevna AKHMADEEVA – she is a junior researcher at the A. P. Ershov Institute of Informatics Systems of Siberian Branch of RAS, an assistant lecturer at the Department of Programming at Novosibirsk State University. List of scientific works includes about 30 publications in the fields of AI, Intelligent System Development and NLP.

email: i.r.akhmadeeva@iis.nsk.su

ORCID: 0000-0002-7371-1087

Материал поступил в редакцию 15 октября 2025 года