

УДК 004.912

doi: 10.15622/rcai.2025.027

## КЛАССИФИКАЦИЯ АРГУМЕНТОВ С ПОМОЩЬЮ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ<sup>1</sup>

Е.А. Сидорова (*lsidorova@iis.nsk.su*)

А.С. Серый (*alexey.seryj@iis.nsk.su*)

И.Р. Ахмадеева (*i.r.akhmadeeva@iis.nsk.su*)

Д.В. Ильина (*dviljina@gmail.com*)

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

Статья посвящена разработке методов автоматической классификации аргументов в русскоязычных текстах с применением генеративных языковых моделей и промпт-инжиниринга. Исследование проводилось на трех датасетах с разметкой аргументации схемами Д. Уолтона: корпусе автоматически сгенерированных англоязычных аргументов NLAS, англоязычных корпусе Aгаucaria и русскоязычном корпусе ArgNetSC. Для классификации аргументов применялись три стратегии: 1) классификация по схемам Уолтона с использованием формальных определений, 2) классификация на основе систематизации схем аргументов, и 3) последовательный вывод с помощью диалога. Исследование с помощью моделей семейства Mistral показало, что наиболее эффективной является диалоговая модель общения с LLM и стратегия автоматического подбора семантически близких примеров для техники Few-Shot. Лучшие оценки составили 0.63 и 0.31  $F_1$ -меры для англоязычных корпусов и 0.15 для русскоязычного корпуса. Для исследования качества предварительной фильтрации классов использовалась модель DeepSeek-R1-Distill-Llama-70B и лучшая стратегия, полученная на первом этапе.  $F_1$ -мера по классам схем составила 0.615, а истинная схема аргумента попала в отфильтрованный список схем в 78.5% случаев.

**Ключевые слова:** анализ аргументации, классификация аргументов, схемы аргументов Уолтона, систематизация схем аргументов, промпт-инжиниринг.

---

<sup>1</sup> Исследование выполнено при финансовой поддержке Российского научного фонда № 23-11-00261, <https://rscf.ru/project/23-11-00261/>.

## Введение

Задача интеллектуального анализа аргументации (Argument Mining, AM) заключается в извлечении аргументов из текста и связывании их в единую структуру, позволяющую проводить дальнейшие исследования, опираясь на цельное представление аргументации заданного текста. Наличие аргументативной структуры позволяет сделать некоторые выводы о характере текста: связности, последовательности рассуждений, наличии или отсутствии обоснования и т.п. Добавление информации о типах используемых аргументативных приемов расширяет область применения AM на более значимые приложения, такие как анализ юридических споров [Walker et al., 2018], [Habernal et al., 2024] и политических дебатов [Lippi et al., 2016], выявление ошибок в рассуждениях и недостоверной информации [Yan et al., 2021], рецензирование проектов [Baimuratov et al., 2025], анализ мнений пользователей и др. Она позволяет оценить, насколько правомерны или убедительны используемые аргументы, поддерживают ли они тезис или опровергают, опираются ли они на факты или на мнения людей.

Существуют различные подходы к моделированию и систематизации аргументов [Lawrence et al., 2016]. Базовой моделью аргумента можно считать модель С. Тулмина [Toulmin, 2003], согласно которой аргумент описывается шестикомпонентой структурой: тезис, посылки, обоснование вывода, поддержка утверждений, границы применимости и степень уверенности. Многие последующие модели аргументации в той или иной степени опираются на эту модель [Freeman, 2011], но на практике имеют более простую структуру. Так, широко известные схемы рассуждения Дугласа Уолтона [Walton, 2009] оперируют только тремя компонентами: тезис, посылки и исключения. Создание онтологии аргументации на основе модели Уолтона [Walton et al., 2008], которая включает около 60 классов (схем аргументации), послужило стимулом многих практических исследований, связанных с разработкой аннотированных ресурсов и ручной разработкой наборов данных (датасетов) для автоматического анализа структурированной аргументации.

Многие исследования [Cabrio et al., 2018], [Zhang et al., 2022], [Shaefer et al., 2021] демонстрируют, что предобученные трансформерные архитектуры достигают высокого качества в различных задачах AM. Однако их эффективность существенно зависит от наличия репрезентативных датасетов. В области AM наблюдается дефицит аннотированных корпусов, а среди существующих лишь немногие содержат детальную классификацию аргументов. В этом контексте появление больших языковых моделей (Large Language Models, LLM) открывает новые перспективы. Благодаря их способности к обобщению языковых паттернов, полученной

в ходе претренинга на разнородных текстах, они потенциально могут решать задачи АМ в условиях few-shot или zero-shot обучения, что является особенно актуальным при недостатке размеченных данных.

Целью данной работы является разработка методов автоматической классификации аргументов в русскоязычных текстах с применением LLM и промпт-инжиниринга. Для достижения поставленных целей в рамках данной работы были сформулированы следующие вопросы исследования.

**RQ1.** Насколько эффективно можно применить LLM и методы промпт-инжиниринга для автоматической многоклассовой классификации аргументов?

**RQ2.** Эффективно ли применение систематизации схем аргументов для предварительного сужения множества классов при многоклассовой классификации аргументов?

Экспериментальное сравнительное исследование проводится на корпусах текстов, снабженных аргументативной разметкой в соответствии с моделью Д. Уолтона.

## 1. Обзор связанных работ

Классификация аргументов – отдельная подзадача АМ, направленная на типизацию уже выявленных аргументативно связанных утверждений согласно заданным категориям. В самом простом случае рассматриваются две категории: поддерживающий и атакующий тип. Классификация аргументов может также включать определение схемы аргумента или его качества (например, сильный аргумент против слабого) или другие признаки, определяемые структурными компонентами аргумента [Wagemans, 2016], [Kononenko et al., 2023].

С развитием средств глубокого обучения нейросетевые подходы стали широко применяться при решении задач АМ, таких как распознавание аргументации в тексте (argument detection) или классификация точки зрения (stance detection) [Ruggeri et al., 2021], Ying et al. 2018]. Более современные исследования все чаще направлены на оценку возможности применения генеративных LLM [Chen et al., 2024]. При этом работ, посвященных проблеме классификации аргументов, значительно меньше, что связано, в том числе, с небольшим количеством размеченных данных. В работе [Pietron et al., 2024] представлена ансамблевая модель, основанная на архитектуре BERT, и ChatGPT-4 в качестве модели тонкой настройки. Модель осуществляла классификацию пар утверждений на трех англоязычных наборах данных, при этом количество классов было невелико – от 2 до 4. Авторы [Bezou-Vrakatseli et al., 2025] проводят сравнительное исследование семи LLM разного размера на задаче Zero-Shot и Few-Shot классификации по семи схемам. При этом утверждается, что данная работа является первым исследованием подобного рода.

Сложность аннотирования богатых аргументами данных, ограничения, связанные с размерами корпусов, разнообразием языков и предметных областей, послужили мотивацией к разработке методики автоматической генерации мультязычного корпуса текстов, размеченных с помощью схем аргументации. Корпус NLAS-multi [Ruiz-Dolz et al., 2024] – мультязычный датасет, аннотированный 20 различными схемами аргументации Д. Уолтона. Корпус создавался с помощью генеративных моделей GPT-3.5-turbo и GPT-4 и ручной фильтрации результатов; соответствие примеров заданной схеме составило более 90%. Хотя этот корпус и является одним из самых больших и богатых корпусов с точки зрения схем аргументации, он состоит из скомпилированных схем, содержащих полные структуры рассуждений, которые редко используются в текстах, написанных людьми.

Качество работы LLM зависит от качества промптов и размера модели. Если выбор модели зависит от финансовых возможностей и ресурсов разработчиков, то выбор стратегии генерации промптов зависит только от них самих.

В работе [Pietron et al., 2024] используется простой запрос вида:

*Это тезис "...". А это утверждение "...". Является ли утверждение аргументом за или против? Напиши ответ одним предложением.*

Анализ ошибок показывает, что модель испытывает трудности при распознавании множественных отрицаний, сарказма, вопросительных утверждений и т.п. Добавление в запрос указания на необходимость объяснить ответ показывает, что модель испытывает трудности и при распознавании логических зависимостей и может противоречить сама себе.

Наша гипотеза заключается в том, что необходимо давать модели больше информации о типах логических связей, разбивать задачу на более простые шаги, приводить примеры (стратегия Few-Shot), использовать диалог для контроля вывода на основе дерева рассуждений.

## **2. Рассуждения на основе систематизации схем аргументов**

Отмечая сложности аннотирования на основе существующей таксономии схем Уолтона, авторы [Сидорова и др., 2024] обосновали необходимость дальнейшей систематизации схем аргументации и предложили их многоаспектную классификацию по четырём основаниям<sup>2</sup>:

**1. Онтологическое отношение.** Включает 11 основных отношений, которыми могут быть связаны пары утверждений.

---

<sup>2</sup> Иерархию и описание схем аргументации можно найти на портале ArgNetBank Studio: <https://uniserv.iis.nsk.su/arg/schemes>.

2. **Направление атаки:** выделяются аргументы, которые могут атаковать тезис, его внешний источник или аргумент в целом, исходя из недостатков оппонента в диалоге. В отдельный класс («Нет атаки») включены аргументы, которые могут быть только поддерживающими.

3. **Тип заключения:** теоретический (вывод носит истинностный характер, сообщает о положении дел или истинности пропозиции) и практический (вывод сообщает о необходимости, долженствовании, возможности или (не)допустимости совершения того или иного действия).

4. **Источник аргумента:** внутренний (вывод сделан логически) или внешний (вывод сделан в силу качеств субъекта).

Таким образом, формально классификатор можно представить как совокупность из четырех ориентированных ациклических графов: , где включает основание , множество классов (промежуточных вершин графа, являющихся дополнительными признаками схем, ), множество схем аргументации (конечные вершины графа, множество общее для всех ) и иерархические связи в графе . Данная структура не является деревом, поскольку схема может соотноситься с несколькими классами (обладать несколькими признаками), но каждый граф обладает свойством: \ является деревом, что можно использовать для последовательного вывода схем сверху-вниз.

Данная категоризация должна позволить генеративной модели последовательно уточнять тип аргумента на основании пересечения его признаков. Взаимодействие с моделью заключается в последовательной классификации аргумента по каждому из оснований классификатора и затем определении его схемы.

Вопросы к модели формируются по следующим принципам. Базовая инструкция ограничивает тематические рамки и сообщает общую задачу модели в диалоге (*Ты помогаешь исследовать приемы аргументации в текстах на русском языке*). Далее следуют формулировка задачи, примеры, список классов и входной аргумент, представляемый в виде пары «**посылка–заключение**». В общем случае аргумент может принадлежать более чем одному классу, поэтому модели ставится задача классификации по многим меткам (Multi-Label). Если семантика классов является взаимоисключающей, это указывается в их описании. Дальнейший диалог продолжается без упоминания входного аргумента и примеров. На каждом шаге модели даются описания классов, являющихся потомками тех классов, которые она назвала на предыдущем шаге, и предлагается уточнить классификацию аргумента.

Для применения генеративных подходов большое значение имеет точность в выборе формулировок для определения каждого классификационного признака. Были применены следующие подходы: (1) выделение ин-

тегрального признака класса; (2) обобщение семантики экземпляров классов; (3) выделение отдельных признаков экземпляров в качестве условия вхождения в класс.

Наиболее корректным представляется использование первого подхода, так как он позволяет выделить семантические признаки аргументативной связи, не представленные в описаниях схем явно. Однако из-за высокой степени обобщенности семантики некоторых классов этот подход не всегда применим. Так, в описании класса «внутренний источник аргумента» реализован второй подход – перечислены возможные условия, при которых пара включается в этот класс: «тезис во фрагменте “заключение” обоснован во фрагменте “посылка” фактами, наблюдениями, логическими рассуждениями, указанием на причинно-следственные, условно-следственные, меронимические, классификационные связи или на последствия действия, о котором идет речь во фрагменте “заключение”».

### 3. Экспериментальное исследование

Для экспериментального исследования использованы три корпуса с разметкой схем аргументации.

1. **Araucaria** [Reed et al., 2008]: 730 аргументов на английском языке, аннотированные 17 схемами Уолтона. Коллекция сильно несбалансированна, для некоторых схем менее пяти примеров («От непоследовательности убеждений», «Опровержение гипотезы», «От исключения», «От страха», «От общей практики»).

2. **NLAS** (англоязычная часть корпуса NLAS-multi) – автоматически синтезированный корпус, содержащий 1893 аргумента на английском языке, аннотированных 20 схемами Уолтона. Корпус NLAS сбалансирован: 75–100 примеров на каждую схему.

3. **ArgNetSC** [Тимофеева и др., 2024] – русскоязычный корпус текстов, относящихся к научной коммуникации, размещенный на платформе ArgNetBank Studio (<https://uniserv.iis.nsk.su/arg>). Содержит более 9 тыс. русскоязычных аргументов, аннотированных 42 схемами Уолтона. Как и Araucaria, коллекция является несбалансированной: 90% аргументов покрывается 17 наиболее частотными схемами.

Основным корпусом, на котором проводились все эксперименты, был ArgNetSC; англоязычные корпуса использовались для получения контрастных метрик для английского языка и синтетических текстов.

#### 3.2. Применяемые подходы

В данной работе исследованы три подхода к решению задачи мультиклассовой классификации аргументов на основе промпт-инжиниринга: (1) классификация по схемам (классам) с использованием определений, заданных в онтологии аргументации, (2) множественная классификация на основе дерева признаков и (3) классификация с помощью диалога с моделью на основе дерева признаков.

1. **Base-classification:** эксперимент, в котором модель сразу решала задачу многоклассовой классификации – определение аргументативной схемы, которой связана заданная пара «посылка–заключение».

2. **Tree-classification:** эксперимент, в котором вначале классификация осуществлялась отдельно по каждому основанию классификатора, а затем по множеству схем, соответствующих найденным классам. Каждый аргумент классифицировался по каждому из классификаторов. На данном этапе модель могла выбирать более одного класса. По результатам классификации вычисляется ограниченное множество схем. На последнем

этапе модель классифицирует аргумент по множеству.

3. **Dialog-classification:** эксперимент в технике Prompt Chaining и в диалоговом режиме, использующий всю иерархию классов схем. Для каждого основания с моделью проводился диалог. На первом шаге требовалось классифицировать  $a$  по множеству. На каждом шаге  $j$  модель уточняла классификацию по множеству, где – результат предыдущего шага.

Диалог завершался по достижении уровня схем. Далее аналогично предыдущему подходу вычислялось множество и проводился финальный этап классификации.

Во всех подходах задача классификации аргументов формулировалась в виде инструкции на естественном языке с применением стратегий Zero-Shot (без добавления примеров) и Few-Shot (с небольшим количеством примеров).

Также задачи разделялись по наличию или отсутствию контекста аргументов в инструкциях (**cont**) и методу подбора примеров в стратегии Few-Shot: случайные примеры (**rand**) или примеры, набираемые из аргументов, близких по смыслу к данному (**sim**). Для сравнения смысловой близости все аргументы были векторизованы моделью LaBSE (<https://huggingface.co/sentence-transformers/LaBSE>).

Результаты усреднялись по пяти случайным разбиениям исходных данных на обучающую и тестовую выборки. При этом каждый аргумент был представлен как минимум в одной тестовой выборке.

### 3.3. Результаты экспериментов

В качестве языковых моделей применялись модели семейства Mistral: Mistral 7b, Mistral NEMO и Mixtral 8x7B. Последняя на текущий момент крупнее большинства моделей, которые можно развернуть в среде с ограниченными вычислительными ресурсами, обладая при этом сравнимой скоростью вывода. Метрики меньших моделей не превысили значений 0.01, поэтому далее показаны только результаты модели Mistral 7x8B.

В табл. 1 представлены результаты классификации аргументов без предварительной фильтрации по классам (приводятся взвешенные  $F_1$ -меры классификации аргументов по схемам).

Таблица 1

Результаты классификации аргументов по схемам

Корпус	Стратегия				
	ZS	ZS+cont	FS-rand	FS-sim	FS-sim+cont
ArgNetSC	0.04	0.04	0.07	<b>0.14</b>	0.12
Araucaria	0.2	0.2	0.22	0.25	0.25
NLAS	0.43	–	0.29	0.63	–

Можно увидеть, что а) случайно выбранные примеры могут привести к увеличению числа ошибок модели даже по сравнению с Zero-Shot, б) правильный выбор примеров дает прирост качества, в) наличие или отсутствие контекста мало влияет на результат. В корпусе NLAS контекст аргументов недоступен, поскольку он содержит только синтетические аргументы. Высокие результаты на корпусе NLAS говорят о том, что искусственно сгенерированные согласно схемам аргументы хорошо распознаются моделью: на большинстве схем модель показала  $F$  0.89–0.99, а общее значение в 0.63 ниже ожидаемого только за счет схем «От знака», «От примера» и «От свидетельских показаний», качество распознавания которых оказалось существенно ниже остальных.

Табл. 2 содержит результаты классификации с учетом классов схем, представленных в п. 2, и предварительной фильтрацией схем (приводятся усредненные взвешенные  $F_1$ -меры и точность фильтрации – доля примеров, когда истинная схема аргумента попала в сокращенный набор схем, полученный путем классификации аргумента по всем основаниям классификатора).

Таблица 2

Результаты классификации аргументов с предварительной фильтрацией по классам схем

Корпус	Стратегия		$F_1$ по классам	Точность фильтрации	$F_1$ по схемам
ArgNetSC	Tree	ZS	0.27	0.06	0.03
		ZS+cont	0.33	0.06	0.04
	Dialog	ZS	0.22	0.1	0.04
		FS-sim	0.55	<b>0.3</b>	<b>0.15</b>
		FS-sim+cont	<b>0.56</b>	<b>0.3</b>	0.14
NLAS	Dialog	FS-sim	0.51	0.37	0.37
Araucaria	Dialog	FS-sim	0.64	0.45	0.31



Как и в первом случае, наличие или отсутствие контекста мало влияет на результат, но правильный выбор примеров дает прирост качества. Для корпуса Arawacaria приведены результаты подхода Dialog как показавшего лучший результат на первом этапе. Определение класса схемы на корпусе NLAS продемонстрировало качество, сопоставимое с результатами, полученными на других корпусах, что привело к снижению качества идентификации схем аргументов.

Качество классификации аргументов, составленных людьми, оказывается значительно ниже. Однако, несмотря на в целом низкое качество определения схемы аргумента, удалось добиться роста среднего качества определения его класса. Таким образом, можно говорить, что генеративная модель без предварительного дообучения показывает низкое качество классификации схемы аргументов, однако такие модели можно применять как помощники при разметке или предварительной фильтрации схем при автоматической классификации аргументов.

Для проверки последней гипотезы был проведен эксперимент по применению самой результативной стратегии (Dialog, FS-sim) в сочетании с моделью большего размера – DeepSeek-R1-Distill-Llama-70B, производной от DeepSeek-R1, способной размышлять последовательно и возвращать ход своих рассуждений; отвечая на вопрос и объясняя причины выбора, она может выступать в качестве ассистента. Модель большего размера, обученная подражать сверхбольшой модели DeepSeek-R1, дала значительный прирост качества:  $F_1$ -мера по классам аргументов составила 0.615, а истинная схема аргумента попала в отфильтрованный список схем в 78,5% случаев. Данный факт позволяет говорить, что генеративные модели могут быть применены в качестве вспомогательного средства при создании аннотированных корпусов аргументов.

Все вычисления выполнялись на следующей конфигурации: AMD Ryzen 9 7950X, 192 GB RAM, NVidia GeForce RTX 4090 24 GB VRAM.

### 3.4. Анализ результатов

Анализ результатов классификации по отдельным схемам показал, что модель наиболее часто ошибочно предсказывает схемы «От вербальной классификации» (398 из 1845), «Существующая практика» (273), «Вербальный скользкий склон» (219) и «Негативные последствия» (129).

Обнаружены устойчивые ошибки, когда модель некорректно интерпретирует примеры одной и той же схемы как примеры других одних и тех же аргументов. Так, 46.7% от всех ошибочных предсказаний схемы «От корреляции к причине» составляют предсказания «От вербальной классификации». Популярность этой схемы может быть связана с тем, что ее большая посылка сформулирована не с указанием семантического отношения, а в общем виде, как *modus ponens*: «Для всех М, если М имеет свойство F, то М можно отнести к имеющим свойство G». Также частотно предсказание «Существующей практики» на месте «От позитивных последствий», «Практического вывода», «От метода» и «От цели к методу».

Это можно объяснить тем, что в каждой из них в одной из посылок эксплицирована причинно-следственная связь между действием и желаемым положением дел.

Значительно снижает общий результат низкий процент распознавания одной из самых частотных схем «От примера». Предполагаемая причина в том, что при отсутствии четких маркеров экзemplификации (*пример, например, в частности*) можно увидеть другие отношения, например каузальности (схема «От причины к следствию») в примере:

*Иногда мы все же заканчиваем работу раньше срока ... В итоге есть куча поводов сорвать срок задачи несмотря на заметный запас подстраховки.*

Отсутствие значительного улучшения при изменении стратегий может быть связано как со слишком общей семантикой классов, так и с несбалансированным подходом к формулированию промптов. Например, описание класса «Гипер-, гипонимия» включает 3 ключевых слова: «пропозиция фрагмента "посылка" является частным случаем, примером пропозиции фрагмента "заключение", и в паре есть пропозиции, связанные семантическим отношением род-вид», а «Каузальность» описана с помощью одного: «пара содержит пропозиции, связанные семантическим причинно-следственным отношением». В примере:

*Экология давно вышла за рамки традиционного толкования, вобрав в себя компоненты других наук. Эти изменения отразились на терминологии отдельных отраслей экологии, которая обогатилась терминами из смежных областей научного знания и понятиями из общей лексики –*

есть указание на оба эти типа отношений, что в целом характерно для текстов со сложным содержанием. Истинным классом здесь является «Каузальность», но модель при любых параметрах возвращает класс «Гипер-, гипонимия» – вероятно потому, что в его описании больше ключевых слов, с которыми соотносится пример.

Кроме того, результат значительно снижает большое количество случаев, когда истинной схемы не оказывается на пересечении классов, что является частотной ситуацией при решении задачи классификации по многим меткам.

## Заключение

В работе представлен подход к автоматической классификации аргументов в соответствии с набором схем Д. Уолтона на основе LLM. Сложность задачи связана как с большим количеством классов, так и с возможностью отнесения схем к нескольким надклассам, что часто приводит к некорректной предварительной фильтрации схем. Рассмотрены три стратегии разработки и применения инструкций: (а) непосредственная классификация аргументов по 48 схемам; (б) использование дополнительной

систематизации схем для вычисления классификационных признаков и предварительной фильтрации множества классов; (в) использование систематизации схем в качестве иерархии для диалогового режима.

Качество классификации аргументов достигает 0.63 F<sub>1</sub>-меры на корпусе автоматически сгенерированных англоязычных аргументов NLAS, 0.31 F<sub>1</sub>-меры на англоязычном корпусе Aгаucia и 0.15 F-меры на русскоязычном корпусе ArgNetSC. Наиболее эффективной стратегией оказался автоматический подбор семантически близких примеров для техники Few-Shot и диалоговая модель общения с LLM. Использование дополнительной систематизации схем аргументов стабильно улучшало показатели на 1.5% на русскоязычных данных и на 6% на англоязычных. Дополнительный эксперимент с моделью большего размера показал, что размер модели существенно влияет на результаты и применение методов дистилляции моделей является перспективным направлением исследований. Можно сделать вывод, что пока генеративные LLM могут использоваться только как вспомогательный инструмент при анализе аргументации.

Возможные методы улучшения результатов включают стандартизацию описаний классов и общую оптимизацию инструкций; реализацию нечеткой фильтрации схем; интеграцию методов на основе SFT и компактных языковых моделей.

### Список литературы

- [Сидорова и др., 2024] Сидорова Е.А., Кононенко И.С. Онтологический анализ приемов аргументации в научном дискурсе // Информационные и математические технологии в науке и управлении. – 2024. – № 3(35). – С. 20-32.
- [Тимофеева и др., 2024] 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. – № 22(1). – С. 27-49.
- [Baimuratov et al., 2025] Baimuratov I., Karpovich A., Lisanyuk E., Prokudin D. Argument Identification for Neuro-Symbolic Dispute Resolution in Scientific Peer Review // In: Proceedings of the 24th ACM/IEEE Joint Conference on Digital Libraries. Association for Computing Machinery, New York, USA, 2025, Article 6. – P. 1-9.
- [Cabrio et al., 2018] Cabrio E., Villata S. Five years of argument mining: a data-driven analysis // In: IJCAI. – 2018. – Vol. 18. – P. 5427-5433. – doi: 10.24963/ijcai.2018/766.
- [Chen et al., 2024] Chen G., Cheng L., Luu A.T., Bing L. Exploring the Potential of Large Language Models in Computational Argumentation // In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics. – 2024. – Vol. 1. – P 2309-2330, Bangkok, Thailand. ACL. – doi: 10.18653/v1/2024.acl-long.126.
- [Freeman, 2011] Freeman J.B. Argument Structure: Representation and Theory. – Springer Science & Business Media, 2011. – Vol. 18.
- [Habernal et al., 2024] Habernal I., Faber D., Recchia N. et al. Mining Legal Arguments in Court Decisions // Artificial Intelligence and Law. – 2024. – No. 32. – P. 1-38.
- [Kononenko et al., 2023] Kononenko I.S., Sery A.S., Shestakov V.K., Sidorova E.A., Zagorulko Y. A. An Approach to Classifying Walton's Argumentation Schemes //

- In: Proc. 2023 IEEE XVI International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE). – 2023. – P. 1540-1545.
- [Lawrence et al., 2016] Lawrence J., Reed C. Argument Mining Using Argumentation Scheme Structures. In: Proceedings of Computational Models of Argument (COM-MA), Potsdam, Germany, 2016. – P. 379-390.
- [Lippi et al., 2016] Lippi M., Torroni P. Argument Mining from Speech: Detecting Claims in Political Debates // In: Proc. Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, Arizona, USA, 2016. – P. 2979-2985.
- [Pietron et al., 2024] Pietron M., Olszowski R., Gomułka J. Efficient argument classification with compact language models and ChatGPT-4 refinements, 2024 // arXiv preprint, <https://arxiv.org/pdf/2403.15473v1>.
- [Reed et al., 2008] Reed C., Mochales Palau R., Rowe G., Moens M.F. Language resources for studying argument // In: Proc. 6th conference on language resources and evaluation (LREC 2008), Marrakech, Morocco, 2008. – P. 91-100.
- [Ruggeri et al., 2021] Ruggeri F., Lippi M., Torroni P. TreeConstrained Graph Neural Networks for Argument Mining. 2021 // ArXiv preprint. doi: 10.48550/arXiv.2110.00124.
- [Ruiz-Dolz et al., 2024] Ruiz-Dolz R., Taverner J., Lawrence J., Reed C. NLAS-multi: A multilingual corpus of automatically // Data in Brief. – 2024. – Vol. 57. – doi: 10.1016/j.dib.2024.111087.
- [Schaefer et al., 2021] Schaefer R., Stede M. Argument mining on twitter: A survey // IT – Information Technology. – 2021. – 63(1). – P. 45-58. – doi: 10.1515/itit-2020-0053.
- [Toulmin, 2003] Toulmin S. The Uses of Argument. – Cambridge University Press, Cambridge, 2003. – 262 p.
- [Wagemans, 2016] Wagemans J.H.M. Constructing a Periodic Table of Arguments // In: Argumentation, Objectivity, and Bias. Proc. of the 11th International Conference of the Ontario Society for the Study of Argumentation (OSSA), 2016. – P. 1-12.
- [Walker et al., 2018] Walker V., Foerster D., Ponce J., Rosen M. Evidence types, credibility factors, and patterns or soft rules for weighing conflicting evidence: Argument mining in the context of legal rules governing evidence assessment // In: Proc. of the 5th Workshop on Argument Mining, Brussels, Belgium, 2018. – P. 68-78.
- [Walton et al., 2008] Walton D., Reed C., Macagno F. Argumentation schemes. – Cambridge University Press, Cambridge, 2008. – 456 p.
- [Walton, 2009] Walton D. Argumentation theory: A Very Short Introduction // In: Argumentation in Artificial Intelligence. – Springer, Boston, 2009. – P. 1-22.
- [Yan et al., 2021] Yan M., Lin Y.R., Litman D. Argumentatively Phony? Detecting Misinformation via Argument Mining. In: Proc. 1st KDD Workshop on AI-enabled Cybersecurity Analytics, 2021.
- [Ying et al. 2018] Ying R., You J., Morris C. et al. Hierarchical graph representation learning with differentiable pooling // ArXiv preprint. – doi: 10.48550/arXiv.1806.08804.
- [Zhang et al., 2022] Zhang G., Nulty P., Lillis D.: A decade of legal argumentation mining: Datasets and approaches // In: Proc. International Conference on Applications of Natural Language to Information Systems. Springer, Cham, 2022. – P. 240-252.