УДК 004.85

# БОЛЬШИЕ ЯЗЫКОВЫЕ МОДЕЛИ В ЗАДАЧЕ РАЗРЕШЕНИЯ ЛЕКСИЧЕСКОЙ МНОГОЗНАЧНОСТИ НА МАТЕРИАЛЕ РУССКОГО ЯЗЫКА

**П. А. Гусяцкая**[1] [0009-0009-2053-8258], **Н. В. Лукашевич**[2] [0000-0002-1883-4121]

[1, 2]*Московский государственный университет имени М. В. Ломоносова, г. Москва, Россия*

[2]*Институт проблем передачи информации имени А. А. Харкевича РАН, г. Москва, Россия*

[1]polinagousyatskaya@gmail.com, [2]louk_nat@mail.ru

***Аннотация***

Статья посвящена описанию результатов экспериментов в области автоматического разрешения неоднозначности (англ. *Word Sense Disambiguation, WSD*) на материале русского языка при помощи генеративных (decoder-only) моделей малого и среднего размеров. Использование генеративных моделей напрямую не является оптимальным подходом к решению данной задачи, однако такие модели имеют потенциал в роли семантических разметчиков необработанных данных. Автоматизация семантической разметки текстов при помощи генеративных моделей потенциально способна преодолеть ограничивающий фактор в виде недостатка размеченных данных для обучения энкодеров.

Как показали более ранние исследования, флагманские англоязычные и мультиязычные модели способны достичь более 90%-ной аккуратности на данной задаче, модели меньшего размера – 80%+. Настоящее исследование ставит своей целью установить, решаема ли аналогичная задача на материале русского языка с помощью русифицированных моделей малого и среднего размеров (до 32 В), не требующих большого количества вычислительных ресурсов для использования.

Эксперименты по разрешению неоднозначности проведены как в базовой постановке (one/few-shot prompting), так и в различных модификациях (обогащение контекста словарной информацией – гиперонимами, гипонимами, метками тематической области и т. д., анализ широкого и узкого контекстных окон неоднозначной лексемы, ансамблевые подходы, в которых одна модель валидирует

и корректирует предсказания другой). В качестве материала исследования использован русскоязычный размеченный ресурс RuSemCor, семантическая разметка которого соответствует категориям семантической сети RuWordNet.

По результатам экспериментов модели показали себя пригодными для решаемой задачи: все модели выходят за уровень случайного предсказания, а наиболее мощные достигают 80%-ной аккуратности, что сопоставимо с результатами англоязычных моделей того же размера. Более информативным для моделей показал себя широкий контекст неоднозначной лексемы. Подходы с дообогащением входных данных и ансамблевые методы дали значительный прирост в качестве.

***Ключевые слова:*** *разрешение лексической неоднозначности, большие языковые модели, компьютерная семантика, классификация.*

**ВВЕДЕНИЕ**

Целью настоящей работы является оценка эффективности русифицированных генеративных моделей малого и среднего размеров в задаче автоматического разрешения неоднозначности (*Word Sense Disambiguation, WSD*). Отметим, что генеративные модели не являются оптимальными для классификационных задач: они подвержены галлюцинациям, испытывают трудности с выводом структурного ответа, а также проигрывают по времени работы. Модели такого типа, однако, обладают двумя весомыми преимуществами. Во-первых, они способны решать различные задачи без предварительного дообучения; во-вторых, они обладают способностью к рассуждению, не реализованной в качестве отдельного модуля логики, а являющейся следствием кодирования информации и ее трансформации на слоях внимания. Все это делает генеративные модели перспективными в качестве семантических разметчиков больших объемов необработанного текста, потенциально способных справиться с недостатком размеченных данных для машинного обучения.

Глобальная цель экспериментов в этой области – это разработка инструмента для семантической разметки русскоязычных текстов, качество разметки которого будет сопоставимо с качеством работы аннотатора-человека.

## БЛИЗКИЕ ПО ТЕМАТИКЕ ИССЛЕДОВАНИЯ

В современной компьютерной лингвистике проблема автоматического разрешения неоднозначности с использованием генеративных моделей находится на начальном этапе проработки. Во многом исследования в этой области зависят от доступности корпусов с семантической разметкой, в особенности, отражающей систему значений многозначных слов. Для английского языка существуют такие ресурсы, как WordNet [1], BabelNet [2] или SyntagNet [3]. Их основная особенность состоит в том, что представление каждого из значений многозначной лексемы содержит множество синтагматических и парадигматических связей, в которые оно вступает. Настоящее исследование основано на аналогичном русскоязычном ресурсе – RuSemCor, семантически размеченном в рамках инвентаря значений русскоязычной семантической сети RuWordNet [4].

Как отмечалось ранее, применение генеративных моделей в решении задачи разрешения лексической многозначности напрямую не является оптимальным подходом. Но при этом генеративные модели рассматриваются как потенциальный источник радикальных, или прорывных, инноваций (англ. *disruptive innovation*), а именно технологий, которые могут обеспечить существенный прорыв в исследованиях в какой-либо области, существенно превосходя по качеству ранее известные методы и подходы [5]. Это и послужило толчком для применения моделей такого типа в том числе к задачам автоматической обработки языка.

Известно несколько фундаментальных исследований генеративных моделей для разрешения лексико-семантической неоднозначности. В работе [5] протестирована зависимость качества определения нужного значения полисеманта от различных техник промптинга: модели предложено или выбрать значение слова из предложенного списка, или оценить истинность суждения типа «в предложении X слово Y употреблено в значении Z». В первом подходе крупные модели (GPT 4) достигают 70% по метрике аккуратности, более легкие модели (Llama-2 7B, 13B) останавливаются на пороге качества случайного предсказания (42.56%, 50.5%). Во втором подходе модели справляются немного лучше – от 53.5% на модели в 7 млрд параметров до 77% на флагманской модели. В указанном исследовании также сформулирована фундаментальная закономерность: качество классификации находится в прямой зависимости от количества параметров модели.

В работе [6] была установлена оптимальная стратегия промптинга для решения задачи разрешения неоднозначности: протестированы техники промпт-инжиниринга, zero-shot- и few-shot-подходы, а также режим эксплицитного последовательного рассуждения, получивший название Chain-of-Thought. Последний подход показал себя как самый результативный для задачи; для дополнительного улучшения качества авторы прибегли к обогащению описаний значений полисемантов синонимами. Совместно с флагманской моделью GPT-4-Turbo подход показал рекордные 90% по метрике аккуратности.

Базовый уровень для исследований на русскоязычном материале намечен в работе [7]. Эксперименты по разрешению многозначности проведены на ряде крупных языковых моделей – как в базовой постановке, так и с применением техник обогащения входных данных словарной информацией. Флагманские англоязычные и мультиязычные модели (GPT-4, Deepseek) достигают качества в 90% и более по метрике аккуратности, модели меньшего размера достигают 80+. Исследования в области разрешения неоднозначности с использованием флагманских генеративных моделей внесли большой вклад в развитие этого направления. Однако, учитывая дороговизну и ресурсозатратность флагманских моделей, такие решения нельзя признать целевыми. Одной из основных задач, таким образом, остается достижение конкуретноспособного качества на моделях меньшего размера.

Эксперименты по проверке информативности широкого и узкого контекстов целевого слова в задаче WSD были выполнены как с помощью ранних подходов, основанных на знаниях [8], так и в нейросетевых подходах (см., например, [9]). Здесь основной исследовательский интерес представляет вопрос, какой контекст содержит больше информативных контекстных маркеров, необходимых для разрешения неоднозначности: глобальный (документ, тематическая область) или локальный (словосочетания). Изучение контекстного окна в сценарии с генеративными моделями имеет значение с точки зрения экономии токенов: если высокая точность достигается на узких контекстах, привлечение широкого контекста оказывается неоправданным, так как будет увеличивать стоимость и время работы модели за счет увеличения промпта без значимого прироста качества.

## НАБОР ДАННЫХ

Материалом исследования послужил корпус русскоязычных текстов RuSemCor, представляющий собой коллекцию текстов, семантически размеченных согласно инвентарю значений русскоязычной семантической сети RuWordNet. Структура ресурса такова: каждой лексеме соответствует одно или несколько значений, определенных по идентификатору синсета и его краткому описанию, а каждому значению – список предложений, в которых реализовано именно это значение.

Эксперименты проведены на части RuSemCor: из корпуса были отобраны неоднозначные лексемы, каждое из значений которых было представлено хотя бы одним предложением. Предложения из корпуса были сгруппированы по неоднозначным лексемам и их значениям. Пример структуры данных представлен на рис. 1. Кроме того, для первоначальных тестов набор данных был урезан до 3000 предложений, поскольку эксперименты с LLM занимают довольно длительное время.


```
"акцент": {
    "107128-N-132865": {
    "RwnSynsetId": "107128-N",
    "RwnSynsetName": "АКЦЕНТ (ПРОИЗНОШЕНИЕ)",
    "sentences": [
        {
        "SentenceId": 66066,
        "Sentence": "Ладно бы еще переозвучивали тех, кто говорит с акцентом - Банионисa, Брыльску,
        Ярвета."
        }
    ]
    },
    "115881-N-132865": {
    "RwnSynsetId": "115881-N",
    "RwnSynsetName": "ПОДЧЕРКНУТЬ, СДЕЛАТЬ АКЦЕНТ",
    "sentences": [
        {
        "SentenceId": 46418,
        "Sentence": "Печально, но pr [в довольно широком смысле] – это та деятельность, которой
        необходимо заниматься регулярно и систематически, причём с акцентом на слово необходимо."
        }
        ]
    }
}
```


Рис. 1. Пример структуры данных из RuSemCor, использованной в экспериментах.

## МОДЕЛИ

Для экспериментов были выбраны две группы генеративных моделей. В первую группу вошли две модели, не превышающие 12 млрд параметров: saiga-llama3 (8b) [10] и Vikhr-Nemo (12B) [11]. В группу моделей среднего размера вошли три модели, насчитывающие от 24 до 32 млрд параметров: модель семейства Vikhr Vistral-24B-Instruct [12], Mistral-Small-24B-Instruct [13] и RuadaptQwen3-32B-Instruct [14]. Все модели доступны на ресурсе HuggingFace через библиотеку transformers.

В экспериментах были использованы только модели типа Instruct, специально обученные на точное выполнение инструкций пользователя. Выбор моделей такого типа обусловлен необходимостью получать от моделей структурный ответ в формате json, чтобы иметь возможность автоматизировать оценку качества классификации. Модели типа Base, работающие по принципу генерации наиболее вероятного следующего токена в цепочке, для классификационных задач, масштабированных на объемные наборы данных, не подходят.

## ОПИСАНИЕ ЭКСПЕРИМЕНТОВ

В рамках проведенных экспериментов на вход моделям были предложены описание задачи и один пример ее решения (англ. one-shot prompting). Кроме базовой постановки задачи были протестированы некоторые подходы к улучшению качества классификации, в частности дообогащение входных данных синтагматической и парадигматической информацией, а также ансамблевый подход, где одна модель валидирует и исправляет ответы другой.


```
{
    "lexeme": "продукт",
    "sentence_id": 47901,
    "sentence_text": "Нарушение правил хранения, закупки или рационального использования зерна и продуктов
    его переработки, а также правил производства продуктов переработки зерна",
    "senses": [
        "106505-N-149796 ПРОДУКТ ТРУДА",
        "106507-N-149796 ПРОДУКТ ПРОИЗВОДСТВА",
        "368-N-149796 ПРОДУКТЫ ПИТАНИЯ",
        "106482-N-149796 РЕЗУЛЬТАТ, ИТОГ"
    ]
}
```


Рис. 2. Входные данные при базовой постановке задачи (полный контекст).

В базовой постановке задачи на вход модели подавалось предложение, содержащее неоднозначную лексему и список ее возможных значений с индексом

значения и кратким описанием, соответствующим наименованию соответствующего синсета в RuWordNet. Пример структуры данных, подаваемой на вход, представлен на рис. 2.

Базовый эксперимент выполнялся с использованием полного контекста неоднозначной лексемы (в среднем 19.5 токенов), а также узкого контекста, когда входное предложение обрезалось до контекстного окна [–2; +2] относительно целевой лексемы, как видно на примере:

(a) "sentence_text": "Пользователи Facebook в скором <TERM>времени</TERM> смогут общаться по Skype через Facebook Connect."

(b) "windowed_context": "в скором <TERM>времени</TERM> смогут общаться"

В подходах с обогащением входных данных краткое описание значения дополнялось списками гипонимов, гиперонимов и меток домена, извлеченных из ресурса RuWordNet для каждого значения целевой лексемы (примеры данных представлены на рис. 3). Модели предлагалось принять решение, обращая внимание на дополнительную словарную информацию. Отметим, что в данном эксперименте в контекст модели привлекается не вся словарная информация, представленная в RuWordNet, а только категории, имеющиеся у большинства лексем. Некоторые категории, такие как синонимы в другой части речи, антонимы, причины, следствия, меронимы или холонимы, у большинства лексем, составляющих RuSemCor, являются пустыми, в связи с чем было принято решение не включать их в эксперимент.

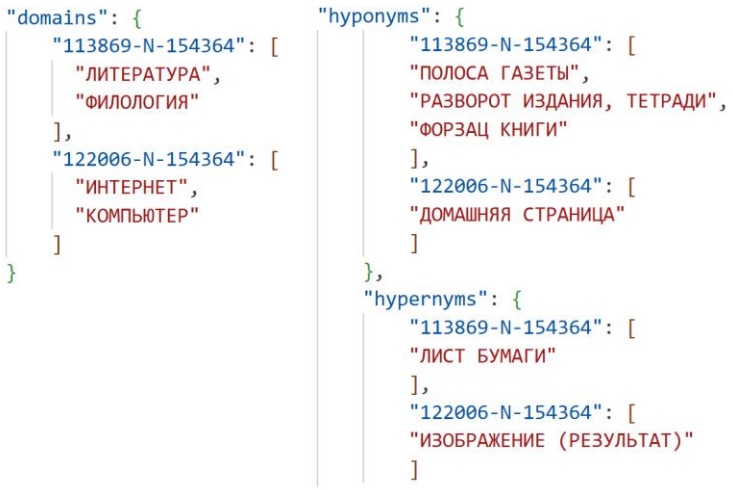


Рис. 3. Пример обогащения лексемы «страница» метками домена (тематической области), гипонимами и гиперонимами.

Ансамблевый подход предполагал следующий сценарий: модель-разметчик проходила по набору данных, выводя в ответ метку предложения, метку нужного значения, а также краткое рассуждение, почему она приняла именно такое решение. Модели-валидатору предлагалось оценить ответы и рассуждения модели-разметчика и исправить их, если решение было оценено как неверное (см. рис. 4).

```
#разметчик
{
    "lexeme": "брать",
    "sentence_id": 108619,
    "sense_id": "115594-V-133880",
    "reason": "Целевое слово используется в контексте ручного действия, что соответствует значению \"брать
    рукой\""
}
#валидатор
{
    "score": 1,
    "reeval_sense_id": null,
    "reason": "Контекст предложения указывает на использование ложки для взятия теста, что соответствует
    значению БРАТЬ РУКОЙ."
}
```

Рис. 4. Пример взаимодействия модели-разметчика и модели-валидатора в ансамблевом подходе.

**РЕЗУЛЬТАТЫ**

Качество классификации оценивалось по метрике accuracy. Результаты моделей по базовому сценарию и сценарию с дообогащением представлены в табл. 1. Как видно по таблице, качество классификации устойчиво растёт с увеличением размера модели: от 0.58 (saiga-llama 8B) до 0.76 (ruadapt 32B) в базовом сценарии. Дообогащение даёт наибольший прирост на крупных моделях (до +7% для vikhr 24B), тогда как для малых моделей эффект минимален (+1%). Сужение контекста, напротив, снижает accuracy на 2–3% для всех моделей.

Результаты ансамблевых подходов (табл. 2) демонстрируют, что ансамблевый подход с более крупным валидатором (mistral 24B) позволяет компенсировать низкое качество малой модели, повышая accuracy saiga-llama 8B с 0.58 до 0.74.

Табл. 1. Результаты классификации.

| | saiga-llama 8B | vikhr-nemo 12B | mistral 24B | vikhr 24B | ruadapt 32B |
|---|---|---|---|---|---|
| Базовый сценарий + широкий контекст | 0.58 | 0.64 | 0.74 | 0.73 | 0.76 |
| Базовый сценарий + узкий контекст | 0.55 | 0.61 | 0.73 | 0.71 | 0.74 |
| Дообогащение | **0.59** | **0.65** | **0.779** | **0.806** | **0.8** |

Табл. 2. Результаты ансамблевых подходов к классификации.

| | Одиночная модель | Ансамблевые подходы | |
|---|---|---|---|
| | saiga-llama 8B | vikhr-nemo12B + saiga_llama 8B | mistral 24B + saiga-llama 8B |
| | acc | acc | acc |
| Базовый сценарий | 0.58 | 0.64 | **0.74** |

Полученные результаты позволили сделать следующие выводы.

1. Задача автоматического разрешения неоднозначности русскоязычных существительных в целом представляется посильной для русифицированных или мультиязычных моделей с количеством параметров больше 8 млрд. Даже модели с 8 млрд параметров показывают качество, превышающее качество случайного выбора, а модели с 24 и больше млрд достигают качества 80%+, приближающегося к отметке, установленной в крупных англоязычных моделях [7].

2. Дообогащение входных данных дополнительной словарной информацией дало прирост качества классификации, в особенности на более крупных моделях, что позволяет признать такой подход целевым для задачи WSD с использованием больших языковых моделей.

3. Короткий контекст показал себя менее информативным для генеративных моделей: по всей видимости, ближайшее окружение неоднозначной лексемы не содержит достаточного количества контекстных маркеров для успешного

разрешения неоднозначности. Глубинный анализ причин таких результатов, однако, требует дальнейшего исследования.

4. Ансамблевые подходы, где модель-валидатор исправляет ответы модели-первичного разметчика, также показали заметный прирост качества. Перспективным сценарием как с точки зрения ресурсов, так и с точки зрения качества является использование более крупной модели в качестве валидатора более легкой. На наших данных ансамблевые методы дали до 16% прироста по метрикам.

## ЗАКЛЮЧЕНИЕ

Проведены эксперименты по автоматическому разрешению неоднозначности русскоязычных лексем с помощью генеративных моделей малого и среднего размеров. Результаты экспериментов показали, что средние русифицированные модели способны осуществлять автоматическую классификацию с уровнем качества 80+, что приближается к целевому уровню, заданному флагманскими моделями. Положительную динамику показали ансамблевые методы классификации, а также подходы с обогащением контекста модели лингвистической информацией. Полученные результаты закладывают основу для разработки масштабного инструмента для автоматической семантической разметки больших объемов текстов.

## Благодарности

Работа выполнена при поддержке гранта РНФ № 24-18-00988.

## СПИСОК ЛИТЕРАТУРЫ

1. *Miller G.* WordNet: A Lexical Database for English // Communications of the ACM. 1995. Vol. 38, No. 11. P. 39–41.

2. *Navigli R., Ponzetto S.P.* BabelNet: Building a very large multilingual semantic network //Proceedings of the 48th annual meeting of the association for computational linguistics. 2010. P. 216–225.

3. *Maru M. et al.* SyntagNet: Challenging supervised word sense disambiguation with lexical-semantic combinations // Proc. of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019. P. 3534–3540.

4. *Loukachevitch N.V., Lashevich G., Gerasimova A.A., Ivanov V.V., Dobrov B.V.* Creating Russian WordNet by Conversion // Proc. of Conference on Computational Linguistics and Intellectual Technologies Dialog-2016, 2016. P. 405–415.

5. *Yae J.H. et al.* Leveraging large language models for word sense disambiguation // Neural Computing and Applications. 2025. Vol. 37, No. 6. P. 4093–4110. https://doi.org/10.1007/s00521-024-10747-5

6. *Sumanathilaka D., Micallef N., Hough J.* Glossgpt: Gpt for word sense disambiguation using few-shot chain-of-thought prompting // Procedia Computer Science. 2025. Vol. 257. P. 785–792. https://doi.org/10.1016/j.procs.2025.03.101

7. *Kirillovich A. et al.* RuSemCor: A Word Sense Disambiguation corpus for Russian //Proc. of the 34th ACM International Conference on Information and Knowledge Management. 2025. P. 6438–6443.
https://doi.org/10.1145/3746252.3761624

8. *Agirre E., Edmonds P. (Eds.*). Word sense disambiguation: Algorithms and applications. Springer Science & Business Media, 2007.
https://doi.org/10.1007/978-1-4020-4809-8

9. *Митрофанова О.А. и др.* Автоматическое разрешение лексико-семантической неоднозначности и выделение конструкций (на материале Национального корпуса русского языка) // Лексикология. Лексикография и Корпусная лингвистика. 2013. С. 122–143.

10. *Gusev I.* Saiga_llama3_8b. Russian Llama-3-based chatbot.
URL: https://huggingface.co/IlyaGusev/saiga_llama3_8b (дата обращения 29.12.2025).

11. Vikhrmodels/Vikhr-Nemo-12B-Instruct-R-21-09-24.
URL: https://huggingface.co/Vikhrmodels/Vikhr-Nemo-12B-Instruct-R-21-09-24 (дата обращения 29.12.2025)

12. Vikhrmodels/Vistral-24B-Instruct.
URL: https://huggingface.co/Vikhrmodels/Vistral-24B-Instruct (дата обращения 29.12.2025)

13. mistralai/Mistral-Small-24B-Instruct-2501.
URL: https://huggingface.co/mistralai/Mistral-Small-24B-Instruct-2501 (дата обращения 29.12.2025)/

14. RefalMachine/RuadaptQwen3-32B-Instruct.

URL: https://huggingface.co/RefalMachine/RuadaptQwen3-32B-Instruct (дата обращения 29.12.2025).

## LLM APPLICATIONS TO THE TASK OF WORD SENSE DISAMBIGUATION IN RUSSIAN TEXTS

**P. A. Gousyatskaya**[1] [0009-0009-2053-8258], **N. V. Loukachevitch**[2] [0000-0002-1883-4121]

[1, 2]*Lomonosov Moscow State University, Moscow, Russia*

[2]*Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Moscow, Russia*

[1]polinagousyatskaya@gmail.com, [2]louk_nat@mail.ru

***Abstract***

This study is dedicated to experiments in Word Sense Disambiguation on Russian-language material using generative (decoder-only) models of small and medium size. While the direct application of generative models is not an optimal approach for this task, such models demonstrate potential as semantic annotators of large volumes of raw data. The automation of semantic text annotation using generative models may help overcome the bottleneck of insufficient labeled data for training encoder-based models.

Previous studies show that state-of-the-art English and multilingual models can achieve accuracy above 90% on this task, while smaller models typically reach 80%+. The present study aims to determine whether a comparable task can be successfully addressed for small- and medium-scale models adapted to Russian, that do not require substantial computational resources.

The experiments reported in this paper are conducted both in a basic setting (one/few-shot prompting) with separate tests for narrow and broad context windows of the ambiguous lexeme, as well as under several modifications, such as context enrichment with paradigmatic information (e.g., hypernyms, hyponyms, domain labels of the target word) and ensemble approaches in which one model validates and refines the predictions of another. The study is based on the Russian annotated resource RuSemCor, annotated in terms of the RuWordNet semantic net.

The experiments ultimately pursue the development of an efficient and accessible pipeline for automatic semantic annotation of Russian texts, useful both for direct application and for preparing annotated data for machinelearning tasks.

***Keywords:*** *word sense disambiguation, LLM, computational semantics, classification.*

**Acknowledgements.** This work was supported by Russian Science Foundation grant No. 24-18-00988.

**REFERENCES**

1. *Miller G.* WordNet: A Lexical Database for English // Communications of the ACM. 1995. Vol. 38, No. 11. P. 39–41.

2. *Navigli R., Ponzetto S.P.* BabelNet: Building a very large multilingual semantic network // Proceedings of the 48th annual meeting of the association for computational linguistics. 2010. P. 216–225.

3. *Maru M. et al.* SyntagNet: Challenging supervised word sense disambiguation with lexical-semantic combinations // Proc. of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019. P. 3534–3540.

4. *Loukachevitch N.V., Lashevich G., Gerasimova A.A., Ivanov V.V., Dobrov B.V.* Creating Russian WordNet by Conversion // Proc. of Conference on Computational Linguistics and Intellectual Technologies Dialog-2016, 2016. P. 405–415.

5. *Yae J.H. et al.* Leveraging large language models for word sense disambiguation // Neural Computing and Applications. 2025. Vol. 37, No. 6. P. 4093–4110. https://doi.org/10.1007/s00521-024-10747-5

6. *Sumanathilaka D., Micallef N., Hough J.* Glossgpt: Gpt for word sense disambiguation using few-shot chain-of-thought prompting //Procedia Computer Science. 2025. Vol. 257. P. 785–792. https://doi.org/10.1016/j.procs.2025.03.101

7. *Kirillovich A. et al.* RuSemCor: A Word Sense Disambiguation corpus for Russian // Proc. of the 34th ACM International Conference on Information and Knowledge Management. 2025. P. 6438–6443.
https://doi.org/10.1145/3746252.3761624

8. *Agirre E., Edmonds P. (Eds.*). Word sense disambiguation: Algorithms and applications. Springer Science & Business Media, 2007.
https://doi.org/10.1007/978-1-4020-4809-8

9. *Mitrofanova O.A. et al.* Avtomaticheskoe razreshenie leksiko-semanticheskoi neodnoznachnosti i vydelenie konstruktsii (na materiale Natsional'nogo korpusa russkogo yazyka) // Leksikologiya. Leksikografiya i Korpusnaya Lingvistika. 2013. P. 122–143.

10. *Gusev I.* Saiga_llama3_8b. Russian Llama-3-based chatbot. URL: https://huggingface.co/IlyaGusev/saiga_llama3_8b (accessed: 29.12.2025).

11. Vikhrmodels/Vikhr-Nemo-12B-Instruct-R-21-09-24. URL: https://huggingface.co/Vikhrmodels/Vikhr-Nemo-12B-Instruct-R-21-09-24 (accessed: 29.12.2025)

12. Vikhrmodels/Vistral-24B-Instruct. URL: https://huggingface.co/Vikhrmodels/Vistral-24B-Instruct (accessed:29.12.2025)

13. mistralai/Mistral-Small-24B-Instruct-2501. URL: https://huggingface.co/mistralai/Mistral-Small-24B-Instruct-2501 (accessed: 29.12.2025)

14. RefalMachine/RuadaptQwen3-32B-Instruct. URL: https://huggingface.co/RefalMachine/RuadaptQwen3-32B-Instruct (accessed: 29.12.2025)

## СВЕДЕНИЯ ОБ АВТОРАХ

***ГУСЯЦКАЯ Полина Андреевна*** — аспирант кафедры теоретической и прикладной лингвистики филологического факультета МГУ. Область научных интересов: автоматическая обработка языка, компьютерная семантика, большие языковые модели, агентные системы.

***Polina A. GOUSYATSKAYA*** is a Ph.D. student at the Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University. Research interests: natural language processing, computational semantics, large language models, agentic systems.

email: polinagousyatskaya@gmail.com

ORCID: 0009-0009-2053-8258

***ЛУКАШЕВИЧ Наталья Валентиновна*** – доктор технических наук, ведущий научный сотрудник НИВЦ МГУ имени М.В. Ломоносова, заведующая кафедрой факультета ВМК, профессор ВМК, филологического факультета МГУ. Область научных интересов: представление знаний, автоматическая обработка текстов, большие языковые модели.

***Natalia V. LOUKACHEVITCH*** is a Doctor of Technical Sciences, leading Researcher at the Research Computing Center of Lomonosov Moscow State University, head of Department at the Faculty of Computational Mathematics and Cybernetics, professor at the Faculty of Computational Mathematics and Cybernetics, Faculty of Philology at Moscow State University. Research interests: knowledge representation, automated text processing, large language models.

email: louk_nat@mail.ru

ORCID: 0000-0002-1883-4121

*Материал поступил в редакцию 3 апреля 2026 года*