

## ФИЛОЛОГИЯ

(шифр научной специальности: 5.9.6)

Научная статья

УДК 81

doi: 10.18522/2070-1403-2023-96-1-59-66

### ИНСТРУМЕНТЫ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ, ОБЛЕГЧАЮЩИЕ ПЕРЕВОДЧЕСКУЮ ДЕЯТЕЛЬНОСТЬ СТУДЕНТОВ НЕЯЗЫКОВЫХ СПЕЦИАЛЬНОСТЕЙ

© Диана Юрьевна Буренкова<sup>1</sup>, Марина Александровна Князева<sup>2</sup>, Мария Сергеевна Невзорова<sup>3</sup>  
<sup>1, 2</sup>Волгоградский институт управления – филиала РАНХиГС, г. Волгоград, Россия; <sup>3</sup>Волгоградский государственный медицинский университет, г. Волгоград, Россия  
<sup>1</sup>dianaymniza2016@gmail.com <sup>2</sup>kravchenko.marina@gmail.com <sup>3</sup>dallas84@mail.ru

**Аннотация.** Раскрываются особенности инструментов компьютерной лингвистики, способные оказать помощь студентам неязыковых факультетов в осуществлении перевода профессионально-ориентированного текста. Рассматриваются общие инструменты: семантические сети, глубокие нейронные сети, – в плане их использования в поиске необходимой информации и для обучения переводу. Показано, что наибольшую популярность среди инструментов компьютерной лингвистики имеют электронные словари, особенно для перевода узкоспециальных текстов. Научная новизна заключается в междисциплинарном рассмотрении вопроса использования машинного перевода в обучении с привлечением трудов современных западных и отечественных ученых и лингвистов. В результате выделены и охарактеризованы основные переводческие платформы, которые могут быть задействованы в обучении студентов неязыковых факультетов. Подчеркивается, что для качественного перевода необходимо задействовать различные инструменты, а также критически оценивать результаты.

**Ключевые слова:** компьютерная лингвистика, вычислительная лингвистика, компьютерные сети, семантические сети, нейронные сети, машинный перевод, электронные словари.

**Для цитирования:** Буренкова Д.Ю., Князева М.А., Невзорова М.С. Инструменты компьютерной лингвистики, облегчающие переводческую деятельность студентов неязыковых специальностей // Гуманитарные и социальные науки. 2023. Т. 96. № 1. С. 59-66. doi: 10.18522/2070-1403-2023-96-1-59-66

## PHILOLOGY

(specialty: 5.9.6)

Original article

### Tools of computational linguistics to facilitate the translation activities of non-language students

© Diana Yu. Burenkova<sup>1</sup>, Marina A. Knyazeva<sup>2</sup>, Maria S. Nevzorova<sup>3</sup>  
<sup>1,2</sup>Volgograd institute of management – Branch of Russian academy of national economy and public administration, Volgograd, Russian Federation; <sup>3</sup>Volgograd state medical university, Volgograd, Russian Federation  
<sup>1</sup>dianaymniza2016@gmail.com <sup>2</sup>kravchenko.marina@gmail.com <sup>3</sup>dallas84@mail.ru

**Abstract.** It is revealed the features of computational linguistics tools that can assist students of non-linguistic faculties in translating professionally oriented text. It is discussed common tools: semantic networks, deep neural networks, in terms of their use in finding the necessary information and for teaching translation. It is shown that electronic dictionaries are the most popular among the tools of computational linguistics, especially for the translation of highly specialized texts. The scientific novelty lies in the interdisciplinary consideration of the use of machine translation in teaching with the involvement of the works of modern Western and domestic scientists and linguists. As a result, the main translation platforms that can be used in teaching students of non-linguistic faculties are identified and characterized. It is emphasized that for high-quality translation it is necessary to use various tools, as well as critically evaluate the results obtained.

**Key words:** computer linguistics, computational linguistics, computer networks, semantic networks, neural networks, machine translation, electronic dictionaries.

**For citation:** Diana Yu. Burenkova, Marina A. Knyazeva, Maria S. Nevzorova Tools of computational linguistics to facilitate the translation activities of non-language students. *The Humanities and Social Sciences*. 2023. Vol. 96. No 1. P. 59-66. doi: 10.18522/2070-1403-2023-96-1-59-66

### *Введение*

На протяжении десятилетий мы наблюдали различные тенденции в развитии вычислительной (компьютерной) лингвистики. Серьезные изменения происходили в мотивах, методах и приложениях. С появлением цифровых технологий лингвистический анализ стал базироваться на новых методах теории информации, которые хорошо сочетались с идеями, доминировавшими в психологии и философии. Затем, вследствие развития когнитивной науки, появилась формальная теория языка и идея прикладной логики. В 50-х гг. прошлого века свое развитие получила прикладная статистика, за которой, в свою очередь, последовала эпоха глубоких сетей. В настоящее время в связи с интенсивным развитием информационных технологий в области компьютерной лингвистики происходят стремительные перемены. Возникает необходимость в исследованиях специалистов-языковедов, которые занимаются не только преподаванием иностранных языков, но и изысканиями в области применения существующих инструментов и технологий компьютерной лингвистики для дальнейшего развития в этой сфере и внедрения этих инструментов в практику преподавания.

### *Обсуждение*

Существуют разные взгляды на то, чем является эта область, и как она должна называться. Определим область вычислительной лингвистики согласно обсуждениям, которые представлены на ведущих площадках, например, таких как Google Scholar.

Название одной из этих площадок, Ассоциации вычислительной лингвистики (ACL), было спорным в 1960-х годах. В настоящее время это название датируется 1968 г. До этого название сообщества включало фразу «Машинный перевод» – тема, которая ранее была более популярна, но сохраняет свою актуальность и до настоящего времени [20].

С тех пор эта область несколько раз меняла направление по ряду причин. Учитывая эту тенденцию к изменениям, а также темпы развития современной науки и вычислительной техники, очевидно, что в будущем их будет еще больше. Вычислительная лингвистика – междисциплинарная область научного знания, которая в разные периоды своего развития тяготела к лингвистике и компьютерным наукам, а в настоящее время наиболее ближе к компьютерному обучению.

Чтобы лучше понимать настоящее и будущее этой области, необходимо знать ее прошлое. Известны основные этапы развития компьютерной лингвистики:

1. (1950 г.): Эмпиризм I – теория информации.
2. (1970 г.): Рационализм I – формальная теория языка и логика.
3. (1990 г.): Эмпиризм II – стохастические грамматики.
4. (2010 г.): Эмпиризм III – глубокие сети.

Через всю историю развития компьютерной лингвистики прослеживаются свои взлеты и падения. Но хорошие идеи редко исчезают полностью, даже в период застоя. Хинтон и Лекун верили в социальные и компьютерные сети и продолжали их использовать, даже когда это было не модно [21; 13; 12; 14]. Солтон работал над проецированием текста в векторные пространства [22; 23; 24; 25; 11].

Исследования машинного перевода, являющегося одной из ключевых тем компьютерной лингвистики, начались со времен появления первых компьютеров и выделились в отдельную область научных исследований после «Джортауновского» эксперимента. С течением времени актуальность подобных исследований только увеличивается в связи с процессом широкой глобализации, общедоступности интернета и стремительным увеличением объема информации. Под влиянием вышеизложенных процессов произошла трансформация языковой подготовки в высшей школе и принятие ФГОС ВО 3++, которые выдвинули на первое место овладение способностью практического применения иностранного языка для профессиональных целей обучающихся на неязыковых факультетах, что в будущем обеспечит выпускникам возможность сохранять конкурентоспособность на рынке. Принимая во внимание тенденцию молодого поколения к использованию машинного перевода, одной из задач преподавателя является необходимость дать выпускнику представление о тех возмож-

ностях, которые дают электронные инструменты для облегчения поиска необходимых информационных ресурсов в научных и профессиональных целях и обучению необходимым для этих целей видам переводческой деятельности.

Рассмотрим общие инструменты компьютерной лингвистики, которые можно использовать для решения поставленных задач.

### 1. Семантические сети

Примерами семантических сетей могут служить WordNet19 [17], CYC [15], Freebase/Wikidata [8] и пр. Многие из этих ресурсов изначально были разработаны для английского языка. Некоторые были расширены для работы с другими языками, так, например, BabelNet 4.021 [18] поддерживает 284 языка. Многие из этих онтологий можно скачать бесплатно. Все они были успешными с точки зрения цитирования.

Масштабы проекта WordNet были ограничены относительно скромным бюджетом. В то время, как Оксфордский словарь английского языка Мюррея [3] и CYC [15] поглощают все доступные ресурсы (и даже больше).

Работа данных ресурсов базируется на моделировании слов (и понятий), которые являются элементами одной схемы (одного графа) и отношения которых представлены как синонимы, антонимы и гипонимы (автомобиль – транспортное средство). Существует множество названий таких графов: графы знаний, семантические сети, онтологии и т.д. Эти графы могут использоваться для представления лексических знаний и/или знаний о мире (факты о мире, выходящие за рамки лингвистики) [16].

Основным отношением между словами в WordNet [17] является синонимия, как, например, связь между словами shut и close или car и automobile. Синонимы – слова, обозначающие одно и то же понятие и взаимозаменяемые во многих контекстах, группируются в неупорядоченные наборы (синсеты). Синсеты состоят из слов, принадлежащих той же части речи, что и исходное слово. Каждый синсет сопровождается небольшой формулировкой, которая разъясняет его значение. Синсеты связаны между собой различными семантическими отношениями.

База данных WordNet [17] подразделяет слова на 117 тыс. синсетов. Многие слова неоднозначны и их относят к двум или более синсетам. Синсеты связаны друг с другом такими отношениями, как гипернимия (is-a) и меронимия (part-whole). Существуют дополнительные связи для глаголов и прилагательных.

В будущем, возможно, такие методы машинного обучения, как заполнение графа знаний, сделают более целесообразным создание таких лингвистических ресурсов, как WordNet [17], CYC [15] и Freebase [8]. Однако, несмотря на впечатляющие результаты, неясно, улучшат ли эти методы охват таких ресурсов, как WordNet и Freebase. Такие отношения, как синонимия, являются (почти) отношениями эквивалентности. Пары слов в одном синсете, скорее всего, будут связаны, а пары слов в разных синсетах вряд ли будут связаны. При заполнении графа знаний, связи распределяются по трем наборам: обучающему, проверочному и тестовому.

### 2. Глубокие (нейронные) сети.

Несомненно, за последние десятилетия глубокие сети добились значительного улучшения производительности. В 2018 году Хинтон, ЛеКун и Бенгио получили премию Тьюринга за это удивительное достижение. В настоящее время большую эффективность показали нейронные сети глубокого обучения. На входной узел нейронной сети подаются любые языковые данные в закодированном виде в виде лексем: буквы, короткие высокочастотные морфемы, биграммы и слова. Применение этого подхода зависит от наличия в распоряжении исследователя большого массива аннотированных текстов.

Модели глубоких сетей, таких как BERT [10], GPT [9], ELMo [19], ERNIE [26], достигли рекордных результатов при решении широкого круга задач. Однако всегда есть возможности для совершенствования.

Наиболее разработанным алгоритмом извлечения именованных сущностей являются системы NER (Name Entity Recognition, NER), т.е. лиц, организаций, географических объектов и т.д. Системы NER, основанные на словарях и правилах, правильно извлекают около 90% сущностей в текстах, в то время как системы на основе BERT уже обеспечивают около 94% правильно извлеченных сущностей [27, с. 85], что можно сопоставить с уровнем человеческой точности и демонстрирует преимущества нейронных сетей глубокого обучения.

В рамках нашего исследования ознакомимся более подробно с некоторыми моделями представления языка. BERT (Bidirectional Encoder Representations from Transformers) в отличие от недавних моделей языковых представлений [19, с. 63] был разработан для предварительного обучения на основе немаркированного текста путем совместного соотнесения левого и правого контекста во всех слоях. В результате предварительно обученная модель BERT может быть точно настроена с помощью всего одного дополнительного выходного слоя для создания современных моделей для широкого круга задач таких, как ответы на вопросы и языковой вывод, без существенных изменений архитектуры, специфичных для конкретной задачи. BERT концептуально прост и эмпирически силен. Он получил новые современные результаты на одиннадцати задачах обработки естественного языка, включая повышение оценки GLUE до 80,5 (абсолютное улучшение на 7,7 пункта), точности MultiNLI до 86,7% (абсолютное улучшение на 4,6%), теста F1 для ответов на вопросы SQuAD v1.1 до 93,2 (абсолютное улучшение на 1,5 пункта) и теста F1 для SQuAD v2.0 до 83,1 (абсолютное улучшение на 5,1 пункта).

Системы нейронного машинного перевода используются основными онлайн-переводчиками (Google Translate, Bing от Microsoft, DeepL, Amazon Translate, Yandex Translate). Однако данные системы все еще имеют много проблем.

Во-первых, остается проблема перевода редких слов и аббревиатур. Зачастую в этих случаях система начинает предлагать новые слова, некорректно переводит имена людей и названия, с которыми не сталкивалась в процессе обучения, или просто не переводит аббревиатуру, оставляя буквы в их исходном значении.

Во-вторых, иногда хороший язык перевода может «расслабить» читателя, и серьезная ошибка в переводе останется незамеченной, что может привести к грубым ошибкам и искажению смысла.

В-третьих, большая часть современных систем переводят предложения в отрыве друг от друга, не учитывая общего контекста. Наиболее успешной разработкой в этой области является онлайн переводчик DeepL, особенно его платные, расширенные версии и, тем не менее, перевод, выполненный в контексте, также нуждается в тщательной проверке и анализе со стороны специалиста-языковеда.

Как показывает практика, самыми востребованными электронными инструментами у студентов для работы с англоязычными текстами, являются онлайн словари.

Таким образом, для перевода узкоспециализированной аутентичной лексики на иностранном языке, целесообразно выделить следующие источники:

– электронные словари, среди которых наибольшей популярностью пользуются Мультитран [4], Оксфордский словарь английского языка Мюррея [3] – средства машинного перевода, которые выполняют на компьютере действия по преобразованию текста для перевода в эквивалентный по содержанию текст на родном языке: онлайн-словарь Prompt [7], Яндекс переводчик [5], DEEPL [6].

Словарь «Мультитран», по мнению Е.Е. Аксеновой, «это даже не словарь, а лексикографический ресурс», словарные значения в котором постоянно пополняются, что является его преимуществом. Он включает в себя многочисленные переводные соответствия из разных источников, в качестве которых может выступать как один автор, так и целая статья. Имена и фамилии авторов переводных соответствий указываются в скобках сразу после них. Сноски на словарь также даются непосредственно сразу после соответствия. И те и другие используются очень активно. Присутствие ссылок на авторство прак-

тически в каждой статье данного словаря является свидетельством его большого прагматического потенциала, что необходимо подчеркнуть при знакомстве студентов с данным инструментом машинного перевода. К недостаткам данного ресурса можно отнести тот факт, что добавить варианты перевода слова может любой желающий, зарегистрировавшийся на сайте, в том числе и дилетант [1, с. 323].

Система Google-Translate и Яндекс МП основаны на статистическом вычислении вероятности совпадений и использует множество баз параллельных текстов, в которых попарно хранятся словосочетания и их переводы. В процессе перевода осуществляется статистический анализ: система подбирает эквивалент для перевода, основываясь на частоте употреблений и подставляет вариант, имеющий наиболее высокий процент совпадений. Однако для каждого языка характерны свои особенности лексико-семантической сочетаемости. Зачастую перевод, выполненный данными программами, свидетельствует о том, что системы не согласовывают слова друг с другом. Обе системы имеют проблемы с правильным переводом терминов, распознаванием частей речи. Встречаются и грамматические ошибки.

DeepL известен как один из самых точных онлайн-переводчиков. Он предлагает бесплатный, простой переводчик на своем сайте, который может переводить текст и файлы на 26 доступных языков. Есть также тарифный план Pro, который открывает неограниченные возможности перевода текста и повышает безопасность данных.

Некоторые из наиболее заметных преимуществ DeepL:

#### 1. Безопасность

Все тексты, использующие DeepL, удаляются сразу после перевода, поэтому сохранность конфиденциальных данных гарантирована. Кроме того, для обеспечения дополнительной безопасности DeepL может производить сквозное шифрование данных, что полностью соответствует политике ЕС по защите данных.

#### 2. Точность

Данный параметр является самым значимым преимуществом данного сервиса. При сравнении с конкурентами DeepL превзошел их в соотношении 3:1, так как работает на основе передовой технологии искусственного интеллекта (AI) и предоставляет множество вариантов настройки перевода. Например, для более естественного и тонкого перевода DeepL предлагает альтернативные слова или фразы, а пользователь выбирает то, что звучит лучше всего. Инструмент даже предлагает на выбор формальные и неформальные голоса в языках, где это очень важно, например, в немецком.

При желании пользователи могут создать свой собственный глоссарий, чтобы определить перевод конкретных слов и фраз в соответствии с их брендом или проектом. Кроме того, DeepL имеет встроенный словарь, позволяющий искать слова и самостоятельно изучать их перевод.

#### 3. Масштабируемость

Все премиум-планы DeepL предлагают неограниченное количество текстовых переводов и позволяют добавлять двух или более пользователей в учетную запись для совместной работы над переводами. Есть возможность переводить ограниченное количество файлов (в зависимости от выбранного вами ценового пакета) размером до 10 МБ, включая HTML-документы. И, наконец, в самых высоких ценовых уровнях DeepL (Advanced и Ultimate) вы можете интегрироваться с вашим CAT-инструментом, таким как Trados Studio, memoQ и Across.

#### *Выводы*

Рассмотрев основные инструменты компьютерной лингвистики, мы видим, что на данный момент не существует идеального словаря, который мог бы с точностью осмыслить все тонкости перевода узкоспециализированных текстов. Поэтому в рамках работы над переводом преподавателям необходимо развивать навыки критического осмысления перевода и умения пользоваться несколькими источниками одновременно.

### Список источников

1. Аксенова Е.Е. Использование ресурсов Интернет при обучении переводу студентов химических специальностей // Информационно-коммуникационные технологии в лингвистике, лингводидактике и межкультурной коммуникации / Под ред. А.Л. Назаренко. М., Университетская книга, 2014. Вып. 6. С. 320–328.
2. Аксенова Е.Е. Обучение письменному переводу в условиях цифровизации высшего образования: вызовы и решения // Языки и культуры в глобальном образовательном пространстве. I Международная научно-практическая конференция. М., 2021. С. 7–11.
3. Оксфордский словарь английского языка Мюррея. – URL: <https://archive.org/details/in.ernet.dli.2015.99992>
4. Онлайн-словарь Мультитран. – URL: <https://www.multitrans.com/>
5. Онлайн-словарь Яндекс переводчик. – URL: [www.translate.yandex.ru](http://www.translate.yandex.ru)
6. Онлайн-словарь DEEPL. – URL: [www.deepl.com/ru/translator](http://www.deepl.com/ru/translator)
7. Онлайн-словарь Promt. – URL: [www.translate.ru](http://www.translate.ru)
8. Bollacker K., Evans C., Paritosh P., Sturge T., Taylor J. Freebase: a collaboratively created graph database for structuring human knowledge // In Proceedings of the ACM SIGMOD International Conference on Management of Data, 2008.
9. Brown T., Mann B., Ryder N. Language models are few-shot learners // In Advances in Neural Information Processing Systems, 2020. P. 1877–1901.
10. Devlin J., Chang M., Lee K., Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding // In NAACL-HLT, 2019.
11. Dubin D. The Most Influential Paper Gerard Salton Never Wrote // Libr. Trends 52, 2004. P. 748–764.
12. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks // International Conference on Neural Information Processing Systems, Lake Tahoe, 3–6 December 2012. P. 1097–1105.
13. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition, 1998.
14. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature, 2015.
15. Lenat D.B. Cyc: A large-scale investment in knowledge infrastructure // Communications of the ACM, 1995. P. 33–38.
16. Mikolov T., Chen K., Corrado G., Dean, J. Efficient Estimation of Word Representations in Vector Space, 2013a.
17. Miller G., Beckwith R., Fellbaum C., Gross D., Miller K.J. Wordnet: An on-line lexical database // International Journal of Lexicography, 1990. P. 235–312.
18. Navigli R., Ponzetto S.P. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network // Artificial Intelligence, 2012. P. 217–250.
19. Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L. Deep contextualized word representations // In NAACL, 2018.
20. Pierce J.R., Carroll J.B. Language and Machines: Computers in Translation and Linguistics, Washington, 1966.
21. Rumelhart D.E, Hinton G.E, Williams R.J. Learning internal representations by error propagation, 1985.
22. Salton G., Fox E.A., Wu H. Extended boolean information retrieval // Communications of the ACM, 1983.

23. *Salton G., Buckley C.* Term-weighting approaches in automatic text retrieval // *Information processing & management*, 1988. P. 513–523.
24. *Salton G., Allan J., Buckley C.* Approaches to passage retrieval in full text information systems // *In ACM SIGIR*, 1993. P. 49–58.
25. *Salton G., Harman D.* *Information Retrieval*. New York, NY: John Wiley & Sons, 2003.
26. *Sun Y., Wang S., Li Y., Feng S., Tian H., Wu H.* Ernie 2.0: A Continual Pre-Training Framework for Language Understanding // *AAAI*, New York, NY, February 2020. P. 7–12.
27. *Wang H.* Robust output feedback stabilization for uncertain discrete-time stochastic neural networks with time-varying delay // *Neural Processing Letters*, 2020. P. 83–103.

### References

1. *Aksenova E.E.* The use of Internet resources in teaching translation to students of chemical specialties // *Information and communication technologies in linguistics, linguodidactics and intercultural communication* / Ed. A.L. Nazarenko. M., University book, 2014. No. 6. P. 320–328.
2. *Aksenova E.E.* Teaching translation in the context of digitalization of higher education: challenges and solutions // *Languages and cultures in the global educational space. I International Scientific and Practical Conference*. M., 2021. P. 7–11.
3. Oxford Murray English Dictionary. – URL: <https://archive.org/details/in.ernet.dli.2015.99992>
4. Multitran Online Dictionary. – URL: <https://www.multitran.com/>
5. Online dictionary Yandex translator. – URL: [www.translate.yandex.ru](http://www.translate.yandex.ru)
6. DEEPL Online Dictionary. – URL: [www.deepl.com/ru/translator](http://www.deepl.com/ru/translator)
7. Online dictionary Promt. – URL: [www.translate.ru](http://www.translate.ru)
8. *Bollacker K., Evans C., Paritosh P., Sturge T., Taylor J.* Freebase: a collaboratively created graph database for structuring human knowledge // *In Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2008.
9. *Brown T., Mann B., Ryder N.* Language models are few-shot learners // *In Advances in Neural Information Processing Systems*, 2020. P. 1877–1901.
10. *Devlin J., Chang M., Lee K., Toutanova K.* BERT: pre-training of deep bidirectional transformers for language understanding // *In NAACL-HLT*, 2019.
11. *Dubin D.* The Most Influential Paper Gerard Salton Never Wrote // *Libr. Trends* 52, 2004. P. 748–764.
12. *Krizhevsky A., Sutskever I., Hinton G.E.* ImageNet Classification with Deep Convolutional Neural Networks // *International Conference on Neural Information Processing Systems*, Lake Tahoe, 3–6 December 2012. P. 1097–1105.
13. *LeCun Y., Bottou L., Bengio Y., Haffner P.* Gradient-based learning applied to document recognition, 1998.
14. *LeCun Y., Bengio Y., Hinton G.* Deep learning // *Nature*, 2015.
15. *Lenat D.B.* Cyc: A large-scale investment in knowledge infrastructure // *Communications of the ACM*, 1995. P. 33–38.
16. *Mikolov T., Chen K., Corrado G., Dean, J.* Efficient Estimation of Word Representations in Vector Space, 2013a.
17. *Miller G., Beckwith R., Fellbaum C., Gross D., Miller K.J.* Wordnet: An on-line lexical database // *International Journal of Lexicography*, 1990. P. 235–312.
18. *Navigli R., Ponzetto S.P.* BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network // *Artificial Intelligence*, 2012. P. 217–250.

19. *Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L.* Deep contextualized word representations // In NAACL, 2018.
20. *Pierce J.R., Carroll J.B.* Language and Machines: Computers in Translation and Linguistics, Washington, 1966.
21. *Rumelhart D.E, Hinton G.E, Williams R.J.* Learning internal representations by error propagation, 1985.
22. *Salton G., Fox E.A., Wu H.* Extended boolean information retrieval // Communications of the ACM, 1983.
23. *Salton G., Buckley C.* Term-weighting approaches in automatic text retrieval // Information processing & management, 1988. P. 513–523.
24. *Salton G., Allan J., Buckley C.* Approaches to passage retrieval in full text information systems // In ACM SIGIR, 1993. P. 49–58.
25. *Salton G., Harman D.* Information Retrieval. New York, NY: John Wiley & Sons, 2003.
26. *Sun Y., Wang S., Li Y., Feng S., Tian H., Wu H.* Ernie 2.0: A Continual Pre-Training Framework for Language Understanding // AAAI, New York, NY, February 2020. P. 7–12.
27. *Wang H.* Robust output feedback stabilization for uncertain discrete-time stochastic neural networks with time-varying delay // Neural Processing Letters, 2020. P. 83–103.

*Статья поступила в редакцию 31.10.2022; одобрена после рецензирования 10.11.2022; принята к публикации 17.11.2022.*

*The article was submitted 31.10.2022; approved after reviewing 10.11.2022; accepted for publication 17.11.2022.*