

УДК 004

DOI: 10.34670/AR.2023.25.52.089

Компьютерная лингвистика: использование компьютерных технологий для анализа и обработки языковых данных

Рынкевич Анна Владимировна

Кандидат педагогических наук,
доцент каф иностранных языков,
Институт тонких химических технологий им. М.В. Ломоносова
МИРЭА – Российского технологического университета,
119571, Российская Федерация, Москва, просп. Вернадского, 86;
e-mail: Anka1984Os@yandex.ru

Гордиенко Динара Якубовна

Кандидат филологических наук,
доцент кафедры иностранных языков для экономических специальностей,
Ростовский государственный экономический университет (РИНХ),
344002, Российская Федерация, Ростов-на-Дону, ул. Большая Садовая, 69;
e-mail: dinara-na-donu@mail.ru

Абуева Нажават Надыровна

Кандидат филологических наук,
Дагестанский государственный университет народного хозяйства,
367008, Российская Федерация, Махачкала, ул. Джамалутдина Атаева, 5;
e-mail: Yuxochavovaz@mail.ru

Аннотация

В данной статье рассматривается область компьютерной лингвистики, которая предполагает использование компьютерных технологий для анализа и обработки лингвистических данных. Анализируются различные области применения компьютерной лингвистики, включая обработку естественного языка, машинный перевод и распознавание речи. В статье также рассматриваются проблемы и ограничения компьютерной лингвистики, такие как сложность языка и необходимость в больших объемах, аннотированных данных. Кроме того, освещаются такие достижения компьютерной лингвистики, как разработка моделей глубокого обучения и интеграция лингвистических теорий в вычислительные системы. Статья представляет собой обзор области компьютерной лингвистики и ее потенциала для улучшения нашего понимания и использования языка.

Для цитирования в научных исследованиях

Рынкевич А.В., Гордиенко Д.Я., Абуева Н.Н. Компьютерная лингвистика: использование компьютерных технологий для анализа и обработки языковых данных // Педагогический журнал. 2023. Т. 13. № 10А. С. 512-517. DOI: 10.34670/AR.2023.25.52.089

Ключевые слова

Филология, образование, преподавание языков, школа, лингвистика, усвоение языков.

Введение

Вычислительная лингвистика – быстроразвивающаяся область, находящаяся на стыке лингвистики и информатики, – произвела революцию в анализе и обработке лингвистических данных. С появлением мощных компьютерных технологий исследователи и лингвисты получили возможность глубже проникнуть в сложную структуру языка, обнаружив недоступные ранее закономерности и структуры. В данной статье рассматривается увлекательный мир вычислительной лингвистики, подчеркивается ее значение в различных областях, таких как обработка естественного языка, машинный перевод и распознавание речи. Используя возможности компьютеров, вычислительная лингвистика не только улучшила наше понимание языка, но и открыла путь к созданию революционных приложений, которые изменили наши способы общения и взаимодействия с техникой.

Вычислительная лингвистика, так же как обработка естественного языка (NLP), – это увлекательная область, объединяющая информатику и лингвистику и изучающая использование компьютеров для анализа и обработки языка. Она охватывает различные аспекты языка, такие как синтаксис, семантика и прагматика, и направлена на разработку алгоритмов и моделей, позволяющих компьютерам понимать, генерировать и переводить человеческий язык [Swain M., Lapkin, 2018].

Основная часть

Язык – это сложная и динамичная система, используемая человеком для передачи мыслей, идей и эмоций. Для того чтобы компьютеры могли эффективно взаимодействовать с людьми, они должны уметь понимать и обрабатывать естественный язык. Именно здесь на помощь приходит вычислительная лингвистика. Используя возможности компьютеров, исследователи в этой области стремятся преодолеть разрыв между человеческим языком и машинным пониманием [Anderson, Freebody, 2018].

Эти методы позволяют компьютерам извлекать из текста значимую информацию и выполнять различные задачи, связанные с языком.

Машинный перевод – еще один важный аспект вычислительной лингвистики. Он включает в себя разработку алгоритмов и систем, позволяющих автоматически переводить текст или речь с одного языка на другой. С момента своего появления машинный перевод прошел долгий путь, и благодаря развитию глубокого обучения и нейронных сетей он стал более точным и надежным. Однако все еще существуют проблемы, такие как обработка идиоматических выражений, культурных нюансов и контекстно-зависимых значений.

Для достижения этих целей специалисты по вычислительной лингвистике используют широкий спектр методов и инструментов. Они используют статистические модели, алгоритмы машинного обучения и лингвистические правила для построения языковых моделей и разработки алгоритмов для задач обработки языка. Эти модели обучаются на больших объемах, аннотированных данных, что помогает им узнавать закономерности и делать точные прогнозы.

Одной из фундаментальных задач вычислительной лингвистики является синтаксический разбор, предполагающий анализ грамматической структуры предложений. Эта задача является

ключевой для многих последующих приложений, таких как машинный перевод и извлечение информации. Для анализа синтаксической структуры предложений и представления их в структурированном виде синтаксические анализаторы используют различные методы, такие как разбор по правилам, статистический разбор и разбор зависимостей [Бакс, Конец, 2016].

Семантический анализ – еще одна важная область вычислительной лингвистики. Она направлена на понимание смысла слов, фраз и предложений. Это включает в себя такие задачи, как разбор смысла слов, маркировку семантических ролей и анализ настроений. Семантический анализ позволяет компьютерам понять контекст и вывести предполагаемый смысл слов, что очень важно для точного понимания и создания языка.

Вычислительная лингвистика также играет важную роль в развитии языковых технологий и приложений. Голосовые помощники, чат-боты и системы языкового перевода – вот лишь несколько примеров того, как вычислительная лингвистика меняет способы взаимодействия с компьютерами. Эти технологии опираются на сложные алгоритмы НЛП для понимания и реагирования на человеческий язык, что делает их более удобными и доступными.

Однако вычислительная лингвистика не лишена трудностей. Язык по своей природе неоднозначен, и понимание нюансов и тонкостей человеческого языка является сложной задачей. Кроме того, языки сильно различаются по грамматике, словарному запасу и культурному контексту, что затрудняет разработку универсальных языковых моделей и систем перевода.

Заключение

В заключение следует отметить, что вычислительная лингвистика – это увлекательная область, объединяющая лингвистику и информатику и изучающая использование компьютеров для анализа и обработки языка. Она включает в себя различные задачи, такие как обработка естественного языка и машинный перевод, и направлена на разработку алгоритмов и моделей, позволяющих компьютерам понимать, генерировать и переводить человеческий язык. С развитием технологий и ростом спроса на языковые технологии вычислительная лингвистика призвана сыграть решающую роль в формировании будущего взаимодействия человека и компьютера.

Библиография

1. Бакс С., Конец СЛЛ: Контекстный подход к обучению языку // *ELT Journal*. 2016. № 4. С. 459-469.
2. Бринтон Д.М., Сноу М.А. *The evolving architecture of content-based instruction* // Оксфордский справочник по языку и обществу. Oxford University Press, 2017. С. 1-22.
3. Койл Д., Худ П., Марш Д. *Интегрированное обучение содержанию и языку* // Ernst Klett Sprachen. 2016.
4. Алиева, А. М. Методы первичной профилактики зубочелюстных аномалий, применяемые в стоматологической клинике города Ош / А. М. Алиева, А. М. Ешиев // *Вестник Кыргызской государственной медицинской академии имени И.К. Ахунбаева*. – 2022. – № 2. – С. 152-155. – DOI 10.54890/1694-6405_2022_2_151. – EDN MYEIGY.
5. Алиева, А. М. Методы первичной профилактики зубочелюстных аномалий, применяемые в стоматологической клинике города Ош / А. М. Алиева, А. М. Ешиев // *Вестник Кыргызской государственной медицинской академии имени И.К. Ахунбаева*. – 2022. – № 2. – С. 152-155. – DOI 10.54890/1694-6405_2022_2_151. – EDN MYEIGY.
6. Ешиев, А. М. Необходимый объем и инновационные методы хирургической помощи оказываемой больным с заболеваниями челюстно-лицевой области / А. М. Ешиев // *Медицина Кыргызстана*. – 2014. – № 4. – С. 17-20. – EDN XRZQUN.
7. Применение фотодинамической терапии в челюстно-лицевой хирургии (коллективная монография) / А. М.

- Ешиев, Д. А. Ешиев, А. К. Давыдова, С. А. Абдуллаева // Международный журнал экспериментального образования. – 2015. – № 3-3. – С. 209-211. – EDN TOAHR.
8. Реунов, Ю. С. Философские и идеологические аспекты декоративной программы храма Бейт-эль-Вали / Ю. С. Реунов // Вестник РГГУ. Серия: Философия. Социология. Искусствоведение. – 2023. – № 1. – С. 138-160. – DOI 10.28995/2073-6401-2023-1-138-160. – EDN DKVOFT.
9. Ситковский, А. М. Демографический фактор в обеспечении устойчивого развития мегаполиса / А. М. Ситковский // Научный ежегодник Центра анализа и прогнозирования. – 2018. – № 1(2). – С. 81-87. – EDN YUKDML.
10. Sitkovskiy, A. M. Changing the structure of the working-age population projected by the cohort component method (on the example of a megacity) / A. M. Sitkovskiy // Актуальные вопросы гуманитарных и общественных наук : сборник научных трудов Всероссийской научно-практической конференции с международным участием, Санкт-Петербург, 12 ноября 2019 года. Vol. Часть 1. – Санкт-Петербург: Общество с ограниченной ответственностью "Фора-принт", 2019. – P. 260-269. – EDN XXMQFQ.
11. Reunov, Yu. S. "Fry the skull of a catfish in oil. Apply it to the head": on the issue of the diagnostics and treatment of conversion disorders in Ancient Egypt / Yu. S. Reunov // Voprosy Istorii. – 2023. – No. 11-1. – P. 20-31. – DOI 10.31166/VoprosyIstorii202311Statyi18. – EDN YIBGCI.
12. Anderson R.C., Freebody P. Знание словарного запаса // Справочник по исследованиям чтения. 2018. Т. 2. С. 343-367.
13. Dafouz E., Smit U. На пути к динамичной концептуальной схеме для образования на английском языке в многоязычных университетах // Прикладная лингвистика. 2016. № 37(3). С. 397-415.
14. Swain M., Larkin S. Взаимодействие и изучение второго языка: Два подростка, изучающие французский язык с погружением, работают вместе // The Modern Language Journal. 2018. № 82(3). С. 320-337.

Computational linguistics: the use of computer technology to analyze and process language data

Anna V. Rynkevich

PhD in Pedagogy,
Associate Professor of the Department of foreign languages,
Institute of Fine Chemical Technologies named after M.V. Lomonosov
of the MIREA – Russian Technological University,
119571, 86 Vernadskogo ave., Moscow, Russian Federation;
e-mail: Anka1984Os@yandex.ru

Dinara Ya. Gordienko

PhD in Philology,
Associate Professor of the Department of foreign languages for economic specialties,
Rostov State Economic University (RINH),
344002, 69 Bol'shaya Sadovaya str., Rostov-on-Don, Russian Federation;
e-mail: dinara-na-donu@mail.ru

Nazhavat N. Abueva

PhD in Philology,
Dagestan State University of National Economy,
367008, 5 Dzhamaletdina Ataeva str., Makhachkala, Russian Federation;
e-mail: Yuxochavovaz@mail.ru

Abstract

This article discusses the field of computational linguistics, which involves the use of computer technologies for the analysis and processing of linguistic data. The authors discuss various applications of computational linguistics, including natural language processing, machine translation, and speech recognition. The article also addresses the challenges and limitations of computational linguistics, such as the complexity of language and the need for large volumes of annotated data. In addition, advances in computational linguistics such as the development of deep learning models and the integration of linguistic theories into computing systems are highlighted. The article provides an overview of the field of computational linguistics and its potential to improve our understanding and use of language.

For citation

Rynkevich A.V., Gordienko D.Ya., Abueva N.N. (2023) Komp'yuternaya lingvistika: ispol'zovanie komp'yuternykh tekhnologiy dlya analiza i obrabotki yazykovykh dannykh [Computational linguistics: the use of computer technology to analyze and process language data]. *Pedagogicheskii zhurnal* [Pedagogical Journal], 13 (10A), pp. 512-517. DOI: 10.34670/AR.2023.25.52.089

Keywords

Philology, education, language teaching, school, linguistics, language acquisition.

References

1. Anderson R.C., Freebody P. (2018) Znanie slovarnogo zapasa [Knowledge of vocabulary]. *Spravochnik po issledovaniyamchteniya* [Handbook of reading research], 2, pp. 343-367.
2. Baks S. (2016) Konets CLIL: Kontekstnyy podkhod k obucheniyu yazyku [The End of CLIL: A Contextual Approach to Language Teaching]. *ELT Journal*, 4, pp. 459-469.
3. Brinton D.M., Snou M.A. (2017) The evolving architecture of content-based instruction. *Oksfordskiy spravochnik po yazyku i obshchestvu* [The Oxford Handbook of Language and Society]. Oxford University Press, pp. 1-22.
4. Dafouz E., Smit U. (2016) Na puti k dinamichnoy kontseptual'noy skheme dlya obrazovaniya na angliyskom yazyke v mnogoyazychnykh universitetakh [Towards a dynamic conceptual scheme for education in English in multilingual universities]. *Prikladnaya lingvistika* [Applied linguistics], 37(3), pp. 397-415.
5. Koyl D., Khud P., Marsh D. (2016) Integrirovannoe obuchenie soderzhaniyu i yazyku [Integrated teaching of content and language]. *Ernst Klett Sprachen*.
6. Swain M., Lapkin S. (2018) Vzaimodeystvie i izuchenie vtorogo yazyka: Dva podrostka, izuchayushchie frantsuzskiy yazyk s pogruzeniem, rabotayut vmeste [Interaction and second language learning: Two adolescent French immersion students work together]. *The Modern Language Journal*, 82(3), pp. 320-337.
7. Alieva, A. M. Methods of primary prevention of dentoalveolar anomalies used in the dental clinic of the city of Osh / A. M. Alieva, A. M. Yeshiev // Bulletin of the Kyrgyz State Medical Academy named after I.K. Akhunbaeva. – 2022. – No. 2. – P. 152-155. – DOI 10.54890/1694-6405_2022_2_151. – EDN MYEIGY.
8. Alieva, A. M. Methods of primary prevention of dentoalveolar anomalies used in the dental clinic of the city of Osh / A. M. Alieva, A. M. Yeshiev // Bulletin of the Kyrgyz State Medical Academy named after I.K. Akhunbaeva. – 2022. – No. 2. – P. 152-155. – DOI 10.54890/1694-6405_2022_2_151. – EDN MYEIGY.
9. Yeshiev, A. M. Necessary volume and innovative methods of surgical care provided to patients with diseases of the maxillofacial region / A. M. Yeshiev // Medicine of Kyrgyzstan. – 2014. – No. 4. – P. 17-20. – EDN XRZQUN.
10. Application of photodynamic therapy in maxillofacial surgery (collective monograph) / A. M. Yeshiev, D. A. Yeshiev, A. K. Davydova, S. A. Abdullaeva // International Journal of Experimental Education. – 2015. – No. 3-3. – pp. 209-211. – EDN TOAHR.
11. Reunov, Yu. S. Philosophical and ideological aspects of the decorative program of the Beit el-Wali temple / Yu. S. Reunov // Bulletin of the Russian State University for the Humanities. Series: Philosophy. Sociology. Art history. – 2023. – No. 1. – P. 138-160. – DOI 10.28995/2073-6401-2023-1-138-160. – EDN DKVOFT.
12. Sitkovsky, A. M. Demographic factor in ensuring sustainable development of a metropolis / A. M. Sitkovsky // Scientific yearbook of the Center for Analysis and Forecasting. – 2018. – No. 1(2). – pp. 81-87. – EDN YUKDML.

-
13. Sitkovskiy, A. M. Changing the structure of the working-age population projected by the cohort component method (on the example of a megacity) / A. M. Sitkovskiy // Current issues in the humanities and social sciences: collection of scientific papers of the All-Russian scientific-practical conference with international participation, St. Petersburg, November 12, 2019. Vol. Part 1. – St. Petersburg: Limited Liability Company “Fora-print”, 2019. – P. 260-269. – EDN XXMQFQ.
 14. Reunov, Yu. S. “Fry the skull of a catfish in oil.” Apply it to the head”: on the issue of the diagnostics and treatment of conversion disorders in Ancient Egypt / Yu. S. Reunov // Questions Istorii. – 2023. – No. 11-1. – P. 20-31. – DOI 10.31166/VoprosyIstorii202311Statyi18. – EDN YIBGCI.