

ПЕДАГОГИКА PEDAGOGICS

УДК 378.147

DOI 10.52575/2712-7451-2023-42-2-233-241

Применение Big Data при обучении английскому языку в высшей школе

Данько Ю.В., Косова Е.И.

Севастопольский государственный университет,
Россия, 299053, г. Севастополь, ул. Университетская, 33
E-mail: bonifacius-zs@yandex.ru, kosova945@mail.ru

Аннотация. На современном этапе обучения иностранному языку возрастает значение цифровых технологий в отборе актуальных дидактических примеров для решения различных задач. Настоящее исследование определяет «большие» данные как основанные на возможностях корпусной лингвистики структурированные массивы данных в обучении лексическим и грамматическим средствам иностранного языка. Целью исследования является систематизация возможностей применения лексических единиц и грамматических структур для обучения английскому языку как общепрофессиональной дисциплине на основе аутентичности, актуальности и завершенности. Возможности «больших» данных в лингводидактике могут способствовать персонификации обучения и стратификации частотности языкового материала по уровням, направленным на приобщение к международному научному и профессиональному дискурсу на любом этапе обучения.

Ключевые слова: «большие» данные, корпус текстов, обучение на основе данных, анализ текстовых данных, цифровые технологии

Для цитирования: Данько Ю.В., Косова Е.И. 2023. Применение Big Data при обучении английскому языку в высшей школе. *Вопросы журналистики, педагогики, языкознания*, 42(2): 233–241. DOI: 10.52575/2712-7451-2023-42-2-233-241

Potentials of Big Data in English Teaching in Higher Education

Yury V. Danko, Ekaterina I. Kosova

Sevastopol State University,
33 Universitetskaya St, Sevastopol 299053, Russia
E-mail: bonifacius-zs@yandex.ru; kosova945@mail.ru

Abstract. The importance of digital technologies in the selection of relevant didactic examples for solving various problems is increasing at the present stage of teaching a foreign language. The present study defines Big Data as structured data arrays based on the possibilities of corpus linguistics in teaching lexical and grammatical means of a foreign language. The goal is to systematize the possibilities of using the corpus of texts in the selection of lexical units and grammatical structures for teaching English as a general professional discipline based on authenticity, relevance and completeness. The possibilities of Big Data in linguodidactics can contribute to the personification of learning and the stratification of the frequency of language material by level, aimed at familiarization with international scientific and professional discourse at any stage of education.

Keywords: Big Data, text corpus, data-based learning, text data analysis, digital technologies



For citation: Danko Yu.V., Kosova E.I. 2023. Potentials of Big Data in English Teaching in Higher Education. *Issues in Journalism, Education, Linguistics*, 42(2): 233–241 (in Russian). DOI: 10.52575/2712-7451-2023-42-2-233-241

Введение

В последние годы «обучение становится адаптивным и личностноориентированным» [Гвозденко и др., 2019, с. 2], что неразрывно связано с цифровизацией образовательного процесса. Хотя в педагогике к настоящему моменту нет общепринятого определения цифровизации образования, главная его черта – применение преподавателем информационно-коммуникационных технологий в профессиональном обучении [Глотова, 2021, с. 1], которые способствуют повышению его качества.

Важнейшими технологиями цифрового (электронного) обучения являются дистанционная, смешанная и гибридная формы [Миленко, 2023, с. 29] с приоритетом смешанной формы обучения [Глотова, 2021, с. 4].

В целом «цифровизация образования представляет собой многогранное явление... меняет не только содержание образования, но и его организацию» [Глотова, 2021, с. 2], и ее примером в отечественном образовании является реализация модели «Цифровой университет», включающей «онлайн-поддержку образовательного процесса, ключевые компетенции цифровой экономики и управление учебным процессом на базе индивидуальной образовательной траектории» [Кузина, 2020, с. 91], ориентированные на «создание, внедрение и совершенствование цифровых сервисов, используемых в сфере науки и высшего образования» [Ращупкина, 2022, с. 3], индивидуальную траекторию, осуществляемую на основе анализа больших данных [Кузина, 2020, с. 93].

Применение различных форм электронного обучения в высшей школе включает анализ массивов структурированных данных о различных субъектах образовательного процесса: студентах, преподавателях, а также о содержании образования (текстовые файлы, аудио и видео материалы), для которого требуются технологии *Big Data* (больших данных) [Нестеренков и др., 2019, с. 243; Глотова, 2021, с. 2]. Дополнение традиционного обучения иностранному языку технологией больших данных является недостаточно исследованным в вопросе содержания коммуникативной компетенции на базовом уровне.

В настоящем исследовании рассмотрены возможности применения больших данных в виде текстовых массивов (возможностей корпусной лингвистики) с целью определения частотной лексики, грамматических структур и их контекста употребления в обучении студентов английскому языку как общепрофессиональной дисциплине.

Объекты и методы исследования

В основу настоящего исследования положены теоретические методы, включающие анализ предшествующих исследований по проблематике цифровизации образовательного процесса [Абдурахманова, Алиев, 2022; Гребенникова, Новикова, 2019; Глотова, 2021; Кудлаев, 2018; Кузина, 2020; Миленко, 2023; Ращупкина, 2022], возможностей корпусной лингвистики в отборе аутентичного дидактического материала по иностранному языку [Антонова, Шангаева, 2020; Балканов, 2022; Митрофанова, Захаров, 2009] и специфики его преподавания в высшей школе [Колкер, Устинова, 2022; Костина, 2016; Новикова, 2022] в контексте возможностей применения технологий больших данных в профессиональном образовании [Белоножко и др., 2017; Гвозденко и др., 2019; Дроздова, 2016; Кондратенко, Кондратенко, 2017; Мамедова и др., 2017; Нестеренков и др., 2019; Шарандин, Булаева, 2023], на основе теоретических положений с учетом актуальных тенденций развития «больших» данных в иноязычной лингводидактике.

Big Data в системе профессионального образования

В широкую практику термин «большие данные» вошел в 2008 году [Дроздова, 2016, с. 2]. К настоящему моменту существует целый ряд его определений, в основе своей включающих совокупность подходов, инструментов, методов и программных технологий, предназначенных для накопления и анализа структурированных и неструктурированных данных с целью извлечения знания, слишком сложных для обработки с использованием традиционных средств [Нестеренков и др., 2019, с. 244 ; Глотова, 2021, с. 3]; в 2017 году в рамках Национальной технологической инициативы (НТИ) они вошли в группу сквозных технологий [Шарандин, Булаева, 2023, с. 225].

Актуальность *Big Data* в последние годы возрастает и в настоящий момент является одним из трендов высшего образования [Шарандин, Булаева, 2023, с. 226], в связи с расширением скорости и доступности Интернета, диджитализацией различных сфер деятельности, появлением облачных хранилищ, цифровых (электронных) сервисов, ростом объемов памяти (в первую очередь жестких дисков, ленточных накопителей и флеш-памяти), что приводит к накоплению информации в виде различных массивов данных.

В контексте развивающейся цифровизации образования «большой объем данных... может быть использован для эффективного управления учебным процессом» [Глотова, 2021, с. 4]; например, применение *Big Data* в профессиональном образовании позволит «преподавателям своевременно получать разнообразную информацию об уровне подготовки студентов, усвоении ими учебных дисциплин, выполнении контрольных заданий и лабораторных работ» [Мамедова и др., 2017, с. 42], контролировать процесс обучения и оперативно вносить изменения [Нестеренков и др., 2019, с. 243], что может быть использовано в организации эффективного образовательного процесса, повышая его доступность и качество [Гвозденко и др., 2019, с. 3]. Поскольку основу технологии составляют данные в цифровом виде, размещенные на физических носителях, возможности ее применения в профессиональном обучении широки и зависят от контекста и условий применения, однако главной целью является расширение возможностей традиционного обучения.

Применительно к обучению выделяются различные виды образовательных данных (educational data): личные (персональные), административные, прогнозные (предполагаемые), о взаимодействии студентов с электронными образовательными ресурсами и друг с другом, об эффективности дидактических материалов [Дроздова, 2016, с. 2].

Применение цифровых данных является частью современного образовательного процесса. Анализ образовательных данных (educational data mining) [Глотова, 2021, с. 2] осуществляется при помощи программных средств и возможностей искусственного интеллекта по следующему алгоритму: «установление цели → определение плана → сбор данных → обработка полученной информации → обеспечение их постоянства → установление закономерностей» [Шарандин, Булаева, 2023, с. 225].

В образовательном процессе анализ данных ориентирован на следующее:

1) повышение эффективности дистанционного обучения: применение различных электронных образовательных ресурсов (ЭОР) накапливает образовательные данные, анализ которых «поможет поднять качество обучения за счет превентивного удовлетворения потребностей в дополнительных учебных материалах» [Нестеренков и др., 2019, с. 245];

2) определение индивидуальной образовательной траектории: на основе множества доступных однотипных моделей обучения можно выстроить индивидуальную образовательную траекторию, персонализированное содержание обучения и выбрать приемлемый способ обучения каждому студенту в развитии необходимых компетенций и довести персональное портфолио на уровень «мета-социальной сети» [Кондратенко, Кондратенко, 2017, с. 114], что может быть определено как «вычислительная педагогика» [Дроздова, 2016, с. 2];

3) выработку эффективной методики обучения дисциплине: на основе анализа данных опыта тысяч преподавателей «методика становится продуктом массового опыта» [Дроздова, 2016, с. 2];



4) эффективную организацию образовательного процесса и формирование необходимых профессиональных навыков: состоит в анализе возможностей более эффективного использования ЭОР, дидактических материалов преподавателями и планировании образовательных интересов студентов [Кондратенко, Кондратенко, 2017, с. 114], интеллектуальных данных и аналитики [Нестеренков и др., 2019, с. 243], связанных с дальнейшей профессиональной ориентацией выпускников и «создание адаптивных систем электронного обучения» [Глотова, 2021, с. 3] в подборе подходящих образовательных программ [Кондратенко, Кондратенко, 2017, с. 114], ориентированных на индивидуализацию обучения и дальнейшее непрерывное самообразование (lifelong-learning) [Глотова, 2021, с. 4] в любом профессиональном аспекте [Белоножко и др., 2017, с. 15];

5) формирование комплекса дидактических упражнений: в контексте обучения иностранным языкам «одним из наиболее эффективных методов обучения иностранному языку является обращение к корпусной лингвистике» [Антонова, Шангараева, 2020, с. 141].

Возможности Big Data в обучении иностранным языкам

В рассмотрении аспектов применения *Big Data* в обучении иностранным языкам прежде всего обратимся к возможностям корпусной лингвистики, дополняющей традиционное обучение, что способствует «успешному овладению новой лексикой и грамматикой, предоставляя возможность опираться на примеры из различных контекстов реальных языковых источников» [Антонова, Шангараева, 2020, с. 142] «на основе аутентичного языкового материала» [Нагель, 2008, с. 55] в решении дидактической проблемы нехватки адекватных педагогических текстовых материалов и актуальных вокабуляров [Нагель, 2008, с. 53] в овладении морфологическими, синтаксическими и лексическими средствами иностранного языка.

Корпус текстов определяется как «набор машиночитаемых, аутентичных текстов, собранных для характеристики или представления состояния, или разнообразия языка» [Антонова, Шангараева, 2020, с. 142], представляющий собой «унифицированный, структурированный, филологически компетентный массив языковых данных» [Кудлаев, 2018, с. 5] «с использованием информационных технологий с программами по обработке текстов» [Абдурахманова, Алиев, 2022, с. 1].

Первый корпус текстов – *Brown Corpus* (объемом 1 миллион слов), представленный в электронном виде в 1960 году Нельсоном Фрэнсисом и Генри Кучерой, был ориентирован на американский вариант английского языка [Антонова, Шангараева, 2020, с. 142] и заложил основу развития корпусной лексикографии. К настоящему времени в связи с расширением возможностей интернет-технологий, объема памяти «облачных» сервисов и электронных ресурсов повышаются возможности работы с текстами и информацией, накоплен большой объем различных текстовых данных и разработаны обширные языковые корпуса (прежде всего в европейских языках).

Применение методов корпусной лингвистики возможно на основе обширного языкового корпуса по отдельному языку, основанному на объемной выборке из различных источников: для английского языка можно выделить *The International Corpus of English*, *The British National Corpus*, *The Freiburg English Dialect Corpus*, *The Bank of English*, *Cambridge and Nottingham Corpus of Discourse in English*, *Michigan Corpus of Academic Spoken English*, *Corpus of Spoken Academic English*, *North America*, *Cambridge International Corpus*; возможен вариант создания частного дидактического корпуса, составленного преподавателем для обучения на конкретных текстах.

Применение в обучении вычислительных методов корпусной лингвистики определяется как «обучение на основе данных» (data-driven learning) [Нагель, 2008, с. 55]. При этом, если первый вид корпуса текстов в первую очередь доступен только онлайн и предполагает обработку данных при помощи специализированных систем поиска языковых выражений в контексте (конкордансеров) [Колкер, Устинова, 2022, с. 3], то в случае со вторым преимуществами являются узкая направленность (специализация) и локальный доступ ввиду

возможности его обработки на любом устройстве, так как такой корпус текстов может быть составлен в любом текстовом редакторе (предпочтительнее в форматах .rtf, .doc, .html, .xml) на основе определенных текстов с возможностью работы непосредственно в самом редакторе, либо возможна обработка массива с помощью программ-анализаторов текста; также его формой может являться применение дидактического параллельного корпуса, содержащего «текст оригинала и его перевод на другой язык» [Антонова, Шангараева, 2020, с. 143].

Поскольку в основе обучения иностранному языку находится овладение определенным уровнем коммуникативной компетенции в совокупности всех навыков, именно отбор статистически значимой лексики для активного и пассивного усвоения является важнейшим преимуществом применения корпуса текстов.

В результате целенаправленного отбора с использованием корпуса текстов формируется «фиксированный набор лексики, которая не потребует большого количества времени для изучения и точно пригодится в процессе общения» [Абдурахманова, Алиев, 2022, с. 2]; при этом объем изучаемого лексического материала устанавливается исходя из различных особенностей обучения иностранному языку.

Big Data в обучении английскому языку в высшей школе

В основе обучения английскому языку как дисциплине базового учебного плана (на языковых факультетах) находится содержательно-ориентированный подход (content-based approach) в совокупности нескольких взаимосвязанных компонентов: лингвистического (лексико-грамматического), риторического (профессионального дискурса в виде частотности тех или иных форм выражения мыслей и их взаимодействия), когнитивного (применения логических задач на английском языке в развитии навыка научной деятельности) и включения английского языка как активатора переработки материала лекции на родном языке [Колкер, Устинова, 2022, с. 2].

Важнейшей задачей обучения английскому языку определяется развитие лингвистического компонента, что неотделимо от изучения иноязычной лексики (активного словарного запаса), ее отбора и актуализации в дидактических упражнениях, способствующих формированию иноязычных навыков в других компонентах.

Применение возможностей корпусной лингвистики при составлении словаря активной лексики английского языка для изучающих получило развитие в 2001 году при создании Macmillan English Dictionary, когда с помощью программ-лемматизаторов были «найжены ряды устойчивых словосочетаний, в результате чего составители... получили экономически эффективный способ автоматического поиска и определения коллокаций для 7 500 наиболее частотных слов английского языка, отобранных для включения в первое издание» [Костина, 2016, с. 106].

В связи с широкими возможностями цифровых технологий в настоящее время «в практической лексикографии наметилась тенденция перехода от традиционных, или бумажных, словарей к словарям электронным» [Pollach, 2012, с. 266], основанным на «отборе данных в корпусе текстов для последующего их включения в словарь» [Балканов, 2022, с. 84], доступных как в оффлайн-, так и в онлайн-формате. Последний является предпочтительным именно в отборе актуальной лексики, сочетаний, клише, и одним из примеров контекстного изучения лексики, основанной на корпусе различных текстов, являются онлайн-словари *Reverso Context* и *Glosbe*, дающие возможность отбора лексем, частотности их употребления «в языковых выражениях на определенных грамматических характеристиках и синтаксических правилах употребления слов» [Митрофанова, Захаров, 2009, с. 322]. Однако применение онлайн-словарей наиболее эффективно в комплексном подходе информационно-образовательной среды всех доступных языковых ресурсов: лингвистических корпусов, подкастов, обучающих программ, веб-приложений [Новикова, 2022, с. 213].



Еще одной возможностью «больших» данных в обучении английскому языку студентов является персонализация обучения [Гвозденко и др., 2019, с. 2] в виде стратификации частотности языкового материала по уровню, направленная на приобщение к международному научному и профессиональному дискурсу на любом этапе обучения [Колкер, Устинова, 2022, с. 1]. Также появляется возможность создания дидактических материалов по английскому языку, учитывающих актуальность, частотность, сочетаемость и контекст употребления лексико-грамматических форм, особенно временных форм английского глагола в области синтаксиса и важнейших глаголов (неправильных, фразовых и вспомогательных) в области морфологии.

Заключение

Определяем *Big Data* перспективным направлением лингводидактики в отборе частотной лексики и контекстного ее употребления, а также овладения основами синтаксиса (в изучении контекстных примеров).

В контексте обучения английскому языку системность отбора языковых средств и их усвоение будут эффективными, если определена цель обучения в рамках избранной специальности, отобраны дидактические текстовые материалы на основе аутентичности, актуальности, завершенности, оформленные в единый текстовый фрагмент (предпочтительнее *.rtf, .doc, .html, .xml*), обрабатываемый при помощи поиска непосредственно в редакторе или с помощью программ анализа текста (в случае офлайн-формата), либо при помощи специализированных систем поиска языковых выражений в контексте (конкордансеров) в случае использования готовых решений с большим объемом выборки.

Полученные результаты отбора лексики, морфологических и синтаксических форм по актуальности, частотности и контекстному употреблению должны отвечать целям обучения, что даст возможность их применения на практических занятиях и поможет в формировании иноязычной коммуникативной компетенции на основе актуальных языковых средств.

Список литературы

- Абдурахманова Л.А., Алиев З.Г. 2022. Использование информационных технологий в процессе изучения английского языка на материале корпусной лингвистики. *Chronos journal*, 7–4 (66): 78–80.
- Антонова Т.Ю., Шангараева Л.Ф. 2020. Корпусная лингвистика и обучение иностранным языкам. *Terra Linguae*, 8: 141–144.
- Балканов И.В. 2022. Корпусная лингвистика: новые возможности лексикографии. *Вестник Дагестанского государственного университета. Серия 2: Гуманитарные науки*, 37(4): 81–86. DOI: [10.21779/2542-0313-2022-37-4-81-86](https://doi.org/10.21779/2542-0313-2022-37-4-81-86)
- Белоножко П.П., Карпенко А.П., Храмов Д.А. 2017. Анализ образовательных данных: направления и перспективы применения. *Науковедение*, 9(4).
- Гвозденко Ю.В., Ищенко А.А., Пилипенко А.В. 2019. Большие данные в системе образования. *Международный студенческий научный вестник*, 5–1: 20.
- Глотова М.И. 2021. Анализ опыта цифровой трансформации отечественного высшего образования. *Современные проблемы науки и образования*, 1: 23. DOI: [10.17513/spno.30503](https://doi.org/10.17513/spno.30503).
- Гребенникова В.М., Новикова Т.В. 2019. К вопросу о цифровизации образования. *Историческая и социально-образовательная мысль*, 11(5): 158–165. DOI: [10.17748/2075-9908-2019-11-5-158-165](https://doi.org/10.17748/2075-9908-2019-11-5-158-165).
- Дроздова А.А. 2016. Перспективы применения технологии «больших» данных в образовании. *Молодой исследователь Дона*, 2(2): 37–39.
- Колкер Я.М., Устинова Е.С. 2022. На пути к внедрению английского языка в профессиональные дисциплины неязыкового профиля. *Иностранные языки в высшей школе*, 1(60): 78–88. DOI: [10.37724/RSU.2022.60.1.008](https://doi.org/10.37724/RSU.2022.60.1.008)
- Кондратенко А.Б., Кондратенко Б.А. 2017. Возможности применения больших данных в образовании в эпоху цифрового общества. *Вестник Калининградского филиала Санкт-Петербургского университета МВД России*, 4(50): 112–115.

- Костина И.А. 2016. Исследование в системе Sketch Engine на примере глаголов физического восприятия в английском языке. *Новая наука: Проблемы и перспективы*, 53(79): 106–111.
- Кудлаев М.С. 2018. Процесс цифровизации образования в России. *Молодой ученый*, 31(217): 3–6.
- Кузина Г.П. 2020. Концепция цифровой трансформации классического университета в «Цифровой университет». *E-Management*, 3(2): 89–96. DOI: [10.26425/2658-3445-2020-2-89-96](https://doi.org/10.26425/2658-3445-2020-2-89-96)
- Мамедова Г.А., Зейналова Л.А., Меликова Р.Т. 2017. Технологии больших данных электронном образовании. *Открытое образование*, 21(6): 41–48. DOI: [10.21686/1818-4243-2017-6-41-48](https://doi.org/10.21686/1818-4243-2017-6-41-48)
- Миленко Н.Н. 2023. Педагогическое проектирование в условиях цифровизации образования. В кн.: Структурная модернизация науки как основа устойчивого развития общества. Сборник статей по итогам Международной научно-практической конференции, Казань, 18 февраля 2023 года. Под ред. А.А. Сукиасяна. Стерлитамак, Уфа, Издательство Агентство международных исследований: 28–37.
- Митрофанова О.А., Захаров В.П. 2009. Автоматизированный анализ терминологии в русскоязычном корпусе текстов по корпусной лингвистике. В кн.: Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог 2009» (Бекасово, 27–31 мая 2009 г.). Выпуск 8(15). Под ред. А.Е. Кибрика. М., РГГУ: 321–328.
- Нагель О.В. 2008. Корпусная лингвистика и ее использование в компьютеризированном языковом обучении. *Язык и культура*, 4: 53–59.
- Нестеренков С.Н., Макаров М.И., Ющенко Н.В., Радкевич А.Д. 2019. Применение больших данных в электронном образовании. В кн.: BIG DATA и анализ высокого уровня. Сборник материалов V Международной научно-практической конференции (Минск, 13–14 марта 2019 г.). В 2 ч. Под ред. В.А. Богуша, В. Zibitsker, С.К. Дика, Д.В. Лихачевского, К.Д. Яшина. Минск, Издательство Белорусского государственного университета информатики и радиоэлектроники. Ч. 2: 242–246.
- Новикова О.В. 2022. Мобильные приложения в педагогическом комплексе обучения английскому языку mobile applications in English language teaching. В кн.: Инновации в профессиональном и профессионально-педагогическом образовании. Материалы 27-й Международной научно-практической конференции (Екатеринбург, 19–20 апреля 2022 г.). Под ред. В.А. Федорова. Екатеринбург, Издательство Российского государственного профессионально-педагогического университета: 210–214.
- Ращупкина А.С. 2022. Стратегия цифровизации высшего образования. *Актуальные вопросы современной экономики*, 6: 809–813.
- Шарандин С.Е., Булаева М.Н. 2023. Технология «большие» данные в профессиональном образовании. В кн.: Профессиональное самоопределение молодежи инновационного региона: проблемы и перспективы. Сборник статей по материалам Всероссийской (национальной) научно-практической конференции (Красноярск, 14–25 ноября 2022 г.). Ч. 2. Красноярск, Челябинск, Нижний Новгород, Москва, Издательство Красноярского государственного аграрного университета: 225–227.
- Pollach I. 2012. Taming Textual Data: The Contribution of Corpus Linguistics to Computer-Aided Text Analysis. *Organizational Research Methods*, 15 (2): 263–287. DOI: [10.1177/1094428111417451](https://doi.org/10.1177/1094428111417451)

References

- Abdurakhmanova L.A., Aliev Z.G. 2022. The use of information technologies in the process of learning english on the material of corpus linguistics. *Chronos journal*, 7–4(66): 78–80 (in Russia).
- Antonova T.Yu., Shangaraeva L.F. 2020. Corpus Linguistics and Foreign Language Teaching. *Terra Linguae*, 8: 141–144 (in Russia).
- Balkanov I.V. 2022. Corpus linguistics: new opportunities of lexicography. *Herald of Dagestan State University*, 37(4): 81–86 (in Russia). DOI: [10.21779/2542-0313-2022-37-4-81-86](https://doi.org/10.21779/2542-0313-2022-37-4-81-86)
- Belonozhko P.P., Karpenko A.P., Khramov D.A. 2017. Algorithm of rangovy optimization of access to data in information system. *Naukovedenie*, 9(4).
- Gvozdenko Yu.V., Ishchenko A.A., Pilipenko A.V. 2019. Big data in the education system. *Mezhdunarodnyy studencheskiy nauchnyy vestnik*, 5–1: 20.
- Glotova M.I. 2021. Analysis of the experience of digital transformation of domestic higher education. *Modern problems of science and education. Surgery*, 1: 23 (in Russia). DOI: [10.17513/spno.30503](https://doi.org/10.17513/spno.30503)
- Grebennikova V.M., Novikova T.V. 2019. On the issue of digitalization of education. *Historical and Social-Educational Idea*, 11(5): 158–165 (in Russia). DOI: [10.17748/2075-9908-2019-11-5-158-165](https://doi.org/10.17748/2075-9908-2019-11-5-158-165).



- Drozdova A.A. 2016. Prospects of "big data" technology application in education. *Young Don Researcher*, 2(2): 37–39 (in Russia).
- Kolker Ya.M., Ustinova E.S. 2022. Towards infiltration of english into academic studies at non-linguistic departments. *Foreign Languages in Tertiary Education*, 1(60): 78–88 (in Russia). DOI: 10.37724/RSU.2022.60.1.008
- Kondratenko A.B., Kondratenko B.A. 2017. The possibilities of using big data in education in the era of digital society. *Bulletin of the Kaliningrad branch of the St. Petersburg University of the Ministry of Interior Affairs of Russia*, 4 (50): 112–115 (in Russia).
- Kostina I.A. 2016. Issledovanie v sisteme Sketch Engine na primere glagolov fizicheskogo vospriyatiya v angliyskom yazyke [Research in the Sketch Engine system on the example of verbs of physical perception in English]. *Novaya nauka: Problemy i perspektivy*, 53(79): 106–111.
- Kudlaev M.S. 2018. Protsess tsifrovizatsii obrazovaniya v Rossii [The process of digitalization of education in Russia]. *Young Scientist*, 31(217): 3–6.
- Kuzina G.P. 2020. The concept of a digital transformation of a traditional university to a “digital university”. *E-Management*, 3 (2): 89–96 (in Russia). DOI: [10.26425/2658-3445-2020-2-89-96](https://doi.org/10.26425/2658-3445-2020-2-89-96)
- Mamedova G.A., Zeynalova L.A., Melikova R.T. 2017. Big data technologies in e-learning. *Open education*, 21(6): 41–48 (in Russia). DOI: [10.21686/1818-4243-2017-6-41-48](https://doi.org/10.21686/1818-4243-2017-6-41-48)
- Milenko N.N. 2023. Pedagogical design in conditions of digitalization of education. In: [Structural Modernization of Science as a Basis for Sustainable Development of Society]. Collection of articles on the results of the International Scientific and Practical Conference (Kazan, February 18, 2023). Ed. A.A. Sukiasyan. Sterlitamak, Ufa, Publ. Izdatel'stvo Agentstvo mezhdunarodnykh issledovaniy: 28–37.
- Mitrofanova O.A., Zakharov V.P. 2009. Avtomatizirovanny analiz terminologii v russkoyazychnom korpusе tekstov po korpusnoy lingvistike [Automated analysis of terminology in the Russian-language corpus of texts in corpus linguistics]. In: *Komp'yuternaya lingvistika i intellektual'nye tekhnologii* [Computational Linguistics and Intelligent Technologies]. Based on the materials of the annual International Conference "Dialogue 2009" (Bekasovo, May 27–31, 2009). Issue 8(15). Ed. A.E. Kibrik. M., Publ. RGGU: 321–328.
- Nagel' O. V. 2008. Corpus linguistics and its use in computer-based language teaching. *Yazyk i kul'tura*, 4: 53–59 (in Russia).
- Nesterenkov S.N., Makarov M.I., Yushchenko N.V., Radkevich A.D. 2019. Primenenie bol'shikh dannykh v elektronnom obrazovanii [Application of big data in e-education]. In: *BIG DATA and Advanced Analytics*. Collection of materials of the V International Scientific and Practical Conference (Minsk, March 13–14, 2019). At 2 pm. Ed. V.A. Bogush, B. Zibitsker, S.K. Dik, D.V. Likhachevsky, K.D. Yashin. Minsk, Publ. Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki. Part 2: 242–246.
- Novikova O.V. 2022. Mobile applications in english language teaching. In: *Innovatsii v professional'nom i professional'no-pedagogicheskom obrazovanii* [Innovations in professional and vocational education]. Proceedings of the 27th International Scientific and Practical Conference (Yekaterinburg, April 19–20, 2022). Ed. V.A. Fedorov. Ekaterinburg, Publ. Rossiyskogo gosudarstvennogo professional'no-pedagogicheskogo universiteta: 210–214.
- Rashchupkina A.S. 2022. Strategy for digital transformation of higher education. *Aktual'nye voprosy sovremennoy ekonomiki*, 6: 809–813 (in Russia).
- Sharandin S.E., Bulaeva M.N. 2023. BIG DATA in professional education. In: *Professional Identity of Youth in Innovative Region*. Materials of the All-Russian (National) Research and Application Conference Problems and Prospects (Krasnoyarsk, November 14–25, 2022). Krasnoyarsk, Publ. Krasnoyarskogo gosudarstvennogo agrarnogo universiteta. Part 2: 225–227 (in Russia).
- Pollach I. 2012. Taming Textual Data: The Contribution of Corpus Linguistics to Computer-Aided Text Analysis. *Organizational Research Methods*, 15(2): 263–287. DOI: [10.1177/1094428111417451](https://doi.org/10.1177/1094428111417451)

Конфликт интересов: о потенциальном конфликте интересов не сообщалось.

Conflict of interest: no potential conflict of interest related to this article was reported.

Поступила в редакцию 15.04.2023

Поступила после рецензирования 10.05.2023

Принята к публикации 10.06.2023

Received April 15, 2023

Revised May 10, 2023

Accepted June 10, 2023



ИНФОРМАЦИЯ ОБ АВТОРАХ

Данько Юрий Владимирович, аспирант кафедры педагогического образования Севастопольского государственного университета, г. Севастополь, Россия

Косова Екатерина Ивановна, ассистент кафедры иностранных языков Севастопольского государственного университета, г. Севастополь, Россия

INFORMATION ABOUT THE AUTHORS

Yury V. Danko, postgraduate student of the pedagogical education department, Sevastopol state university, Sevastopol, Russia

Ekaterina I. Kosova, assistant of foreign languages department, Sevastopol state university, Sevastopol, Russia