

Визуализация данных динамики зарплат и спроса на AI-специалистов по странам мира с использованием библиотеки Seaborn

Федосеев Константин Андреевич

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

В статье представлено практическое применение библиотеки Seaborn языка Python для визуализации и анализа динамики мирового рынка специалистов в области искусственного интеллекта. На основе датасета, содержащего информацию о вакансиях, уровне зарплат, проценте удаленной работы и востребованных навыках в семи странах за период 2020-2026 годов, построена система взаимодополняющих графиков: линейные диаграммы для отслеживания временных трендов, тепловые карты для пространственного анализа, боксплоты для сравнения статистических распределений и точечный график для отслеживания изменений. Результаты визуализации позволили идентифицировать ключевые закономерности: географическую дифференциацию уровня оплаты труда, цикличность спроса на AI-специалистов, растущую значимость навыков работы с Python и PyTorch, а также отсутствие прямой зависимости между процентом удаленной работы и размером заработной платы. Представленная методология может быть использована для мониторинга конъюнктуры рынка труда в различных отраслях и регионах.

Ключевые слова: визуализация данных, искусственный интеллект, AI-специалисты, рынок труда, Python программирование, библиотека Seaborn, анализ зарплат, временные ряды, геоаналитика, навыки программирования, удаленная работа, дата-аналитика, тепловые карты, сравнительный анализ.

Visualization of salary dynamics and demand for AI specialists by country using the Seaborn library

Fedoseev Konstantin Andreevich

Sholom-Aleichem Priamursky State University

Student

Abstract

The article presents a practical application of the Seaborn library in Python for visualizing and analyzing the dynamics of the global market for artificial intelligence specialists. Based on a dataset containing information about vacancies, salary levels, remote work percentages, and in-demand skills in seven countries for the period 2020-2026, a system of complementary graphs was constructed: line charts for tracking temporal trends, heat maps for spatial analysis, box plots for comparing

statistical distributions, and a scatter plot for tracking changes. The visualization results made it possible to identify key patterns: geographical differentiation of wage levels, cyclicity of demand for AI specialists, the growing importance of Python and PyTorch skills, and the absence of a direct relationship between the percentage of remote work and salary size. The presented methodology can be used to monitor the labor market situation in various industries and regions.

Keywords: data visualization, artificial intelligence, AI specialists, labor market, Python programming, Seaborn library, salary analysis, time series, geospatial analysis, programming skills, remote work, data analytics, heat maps, comparative analysis.

1 Введение

Актуальность

Актуальность визуализации данных в современном информационном обществе определяется экспоненциальным ростом объемов генерируемой информации, которую невозможно эффективно анализировать традиционными методами. Визуализация преобразует сложные многомерные массивы данных в интуитивно понятные графические образы, позволяя выявлять скрытые закономерности, аномалии и тренды, неочевидные при табличном представлении. В условиях цифровой трансформации всех сфер деятельности визуальные методы анализа становятся критически важными для принятия обоснованных решений на основе данных (data-driven decision making). Особую значимость приобретает интерактивная визуализация, обеспечивающая глубокое погружение в данные через drill-down анализ и многомерное исследование взаимосвязей. Развитие специализированных библиотек, таких как Seaborn, Matplotlib и Plotly, демократизирует доступ к сложным методам визуализации, позволяя исследователям и аналитикам без глубоких знаний в компьютерной графике создавать публикационно-качественные графики.

Обзор исследования

Методологическую основу формируют работы, посвященные применению языка Python: О.К. Фарайдунов и З.А. Амонбекова [2] обосновывают его использование для решения экономических задач, а сравнительный анализ А. Н. Sial и соавторов [6] выделяет библиотеку Seaborn как оптимальный инструмент для статистической визуализации, что подтверждается практическим руководством VanderPlas [8]. Предметную область определяют исследования Э.С. Шарапова и Д.С. Корневой [3], непосредственно анализирующие датасеты зарплат в Data Science, а также работы Н.А. Олимова [4] и Е.А. Тарана [5], раскрывающие контекст цифровой экономики и структурные изменения цифрового капитализма, в рамках которых функционирует рынок AI-специалистов.

Глобальный контекст и прогнозные оценки, необходимые для интерпретации результатов, задаются авторитетными источниками. Отчет World Economic Forum [9] фиксирует растущий мировой спрос на

специалистов в области данных и искусственного интеллекта, подтверждая актуальность темы. Исследование Р. Rajpurkar и М. Р. Lungren [7], рассматривающее будущее ИИ в медицине, служит примером углубленного отраслевого прогноза и демонстрирует методологию анализа долгосрочных технологических трендов, аналогии которой применяются в данном исследовании для рынка труда.

ель исследования

Цель исследования – разработка методологии визуализации данных динамики зарплат и спроса на AI-специалистов с использованием библиотеки Seaborn для выявления закономерностей развития глобального рынка труда.

2 Материалы и методы

В качестве основного источника данных для визуализации и анализа использован файл `country_ai_trends.csv` в составе датасета «Global AI and Data Science Job Market (2020-2026)», размещенного на платформе Kaggle [1]. Данный файл содержит агрегированные статистические показатели рынка труда в сфере искусственного интеллекта и Data Science по странам за период с 2020 по 2026 год. В качестве инструментария использовался язык Python 3.14 с библиотеками Pandas для обработки данных, NumPy для математических операций, и Seaborn как основной инструмент статистической визуализации в сочетании с Matplotlib. Применены методы дескриптивной статистики, сравнительного и корреляционного анализа. Разработана система взаимодополняющих визуализаций: линейные диаграммы для временных трендов, тепловые карты для пространственного анализа зарплат, боксплоты для сравнения статистических распределений, точечные диаграммы для исследования взаимосвязей и столбчатые графики для анализа категориальных данных. Методология включает этапы подготовки данных, анализа, разработки визуализаций и интерпретации результатов.

3 Результаты и обсуждение

В результате выполнения программы сформирована комплексная система визуализации данных о рынке AI-специалистов. Программа последовательно генерирует три отдельных окна с графиками и предоставляет текстовую аналитическую сводку в консоли.

Первое окно демонстрирует основные аспекты анализа. Линейный график (рис. 1) выявляет выраженную динамику количества AI-вакансий по странам. США и Индия показывают наибольшие абсолютные значения, однако США демонстрируют значительные колебания между годами, в то время как Индия сохраняет относительно стабильный рост. Германия и Канада имеют более сглаженные траектории, что указывает на устойчивость их рынков.



Рис. 1. Линейчатый график

Тепловая карта (рис. 2) визуализирует географическую и временную дифференциацию средних зарплат в долларах США (USD). Наиболее интенсивные цвета (высокие зарплаты) сосредотачиваются в Канаде (2026 г.), США (2022, 2025 гг.) и Индии (2021 г.). Наблюдается пространственная кластеризация: англоязычные страны (США, Канада, Австралия, Великобритания) формируют зону с повышенным уровнем оплаты, хотя внутри нее существуют значительные различия по годам.

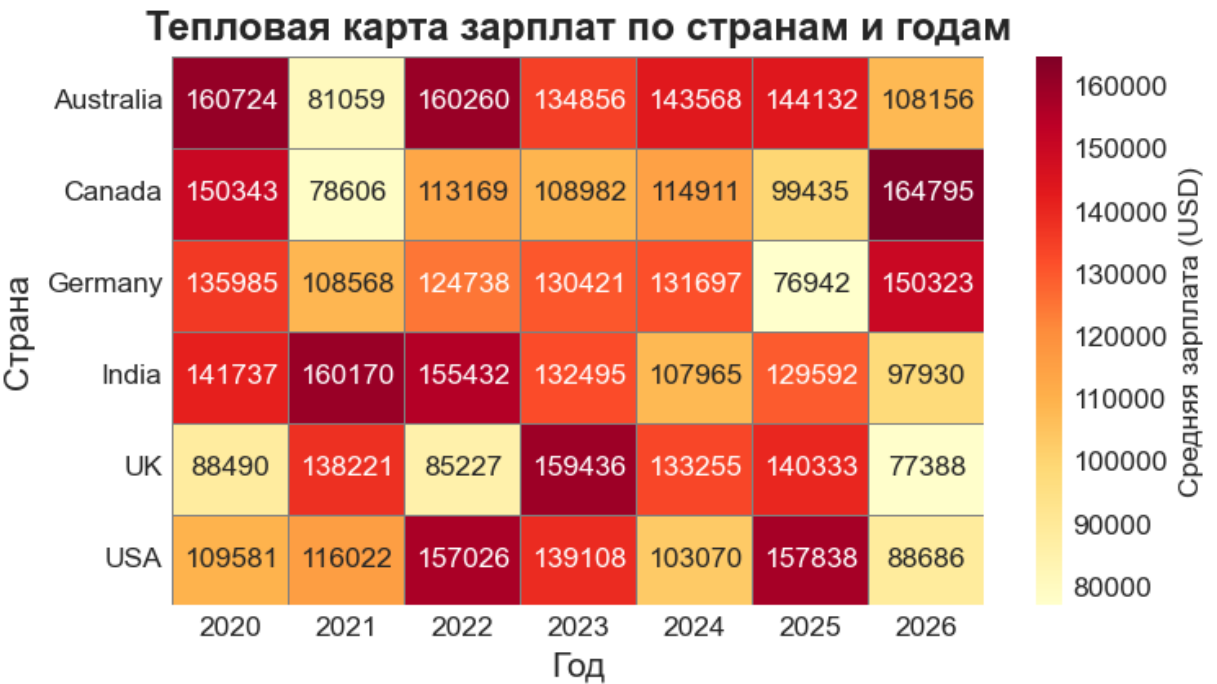


Рис. 2. Тепловая карта

Столбчатая диаграмма (рис. 3) распределения топ-навыков за 2026 год определяет Python и PyTorch как доминирующие требования рынка. Их популярность охватывает большинство анализируемых стран. Навыки работы с облачными платформами (AWS, GCP, Azure) также сохраняют высокую

востребованность, что отражает общую тенденцию к облачной инфраструктуре в AI-проектах.

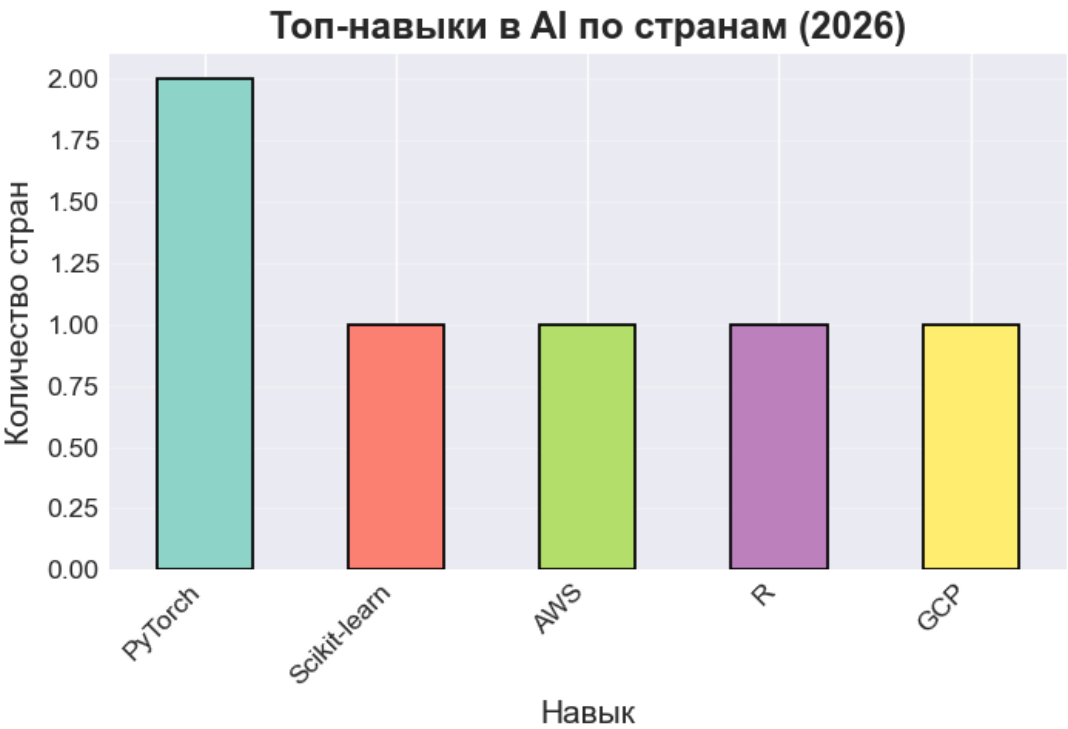


Рис. 3. Столбчатая диаграмма

Точечная диаграмма (рис. 4) исследует взаимосвязь между процентом удаленной работы и уровнем зарплаты. Результаты визуализации не обнаруживают четкой линейной зависимости. Высокие зарплаты встречаются как при высоком (более 70%), так и при низком (менее 30%) проценте удаленной работы. Это указывает на то, что формат занятости не служит ключевым фактором при определении размера компенсации для AI-специалистов.

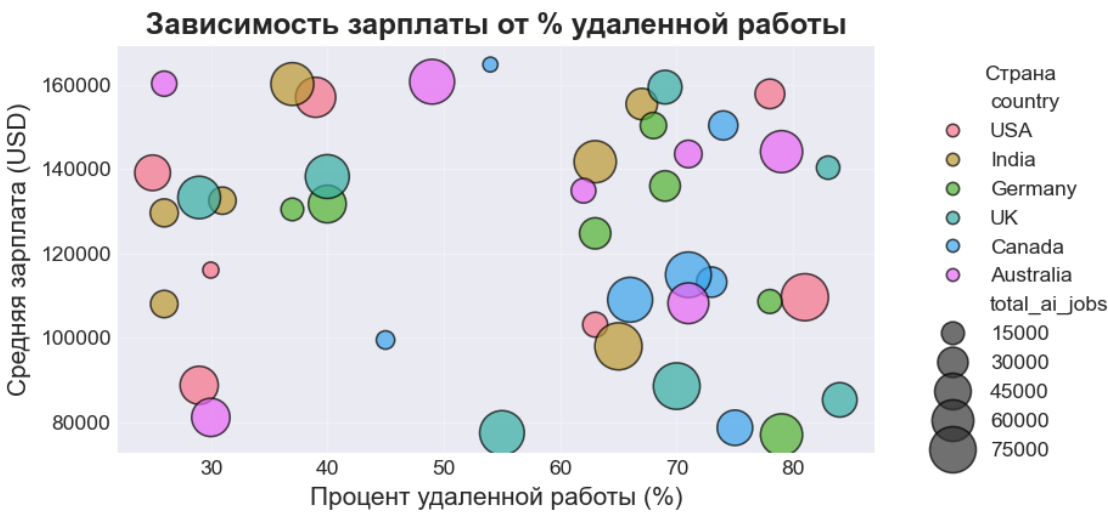


Рис. 4. Точечная диаграмма

Второе окно с боксплотом (рис. 5) представляет статистическое распределение зарплат по странам за весь период. Боксплоты четко визуализируют размах данных, положение медианы и наличие выбросов. Канада показывает самый высокий выброс (2026 г.), что соответствует максимальной зарплате в датасете. Распределение зарплат в США имеет наибольший межквартильный размах, подтверждая высокую волатильность рынка. Индия, при высоких медианных значениях, демонстрирует компактное распределение, что свидетельствует о меньшем разбросе зарплат внутри страны.

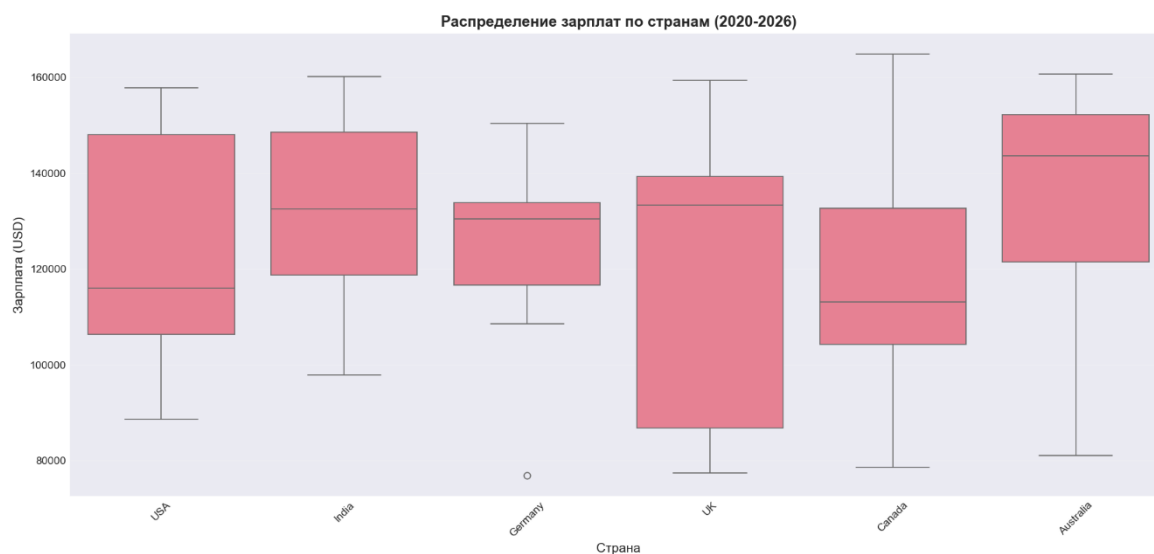


Рис. 5. Боксплот

Третье окно (рис. 6) содержит сетку из шести совмещенных графиков, каждый из которых посвящен отдельной стране. На графиках две оси Y: левая (синяя линия с маркерами-кругами) отображает динамику количества вакансий, правая (красная линия с маркерами-квадратами) – динамику средней зарплаты. Данная визуализация позволяет проводить парный анализ двух ключевых метрик в их временном развитии. Например, для США в 2022 году наблюдается синхронный пик по обоим показателям, тогда как в Индии пик зарплат (2021 год) опережает пик количества вакансий.

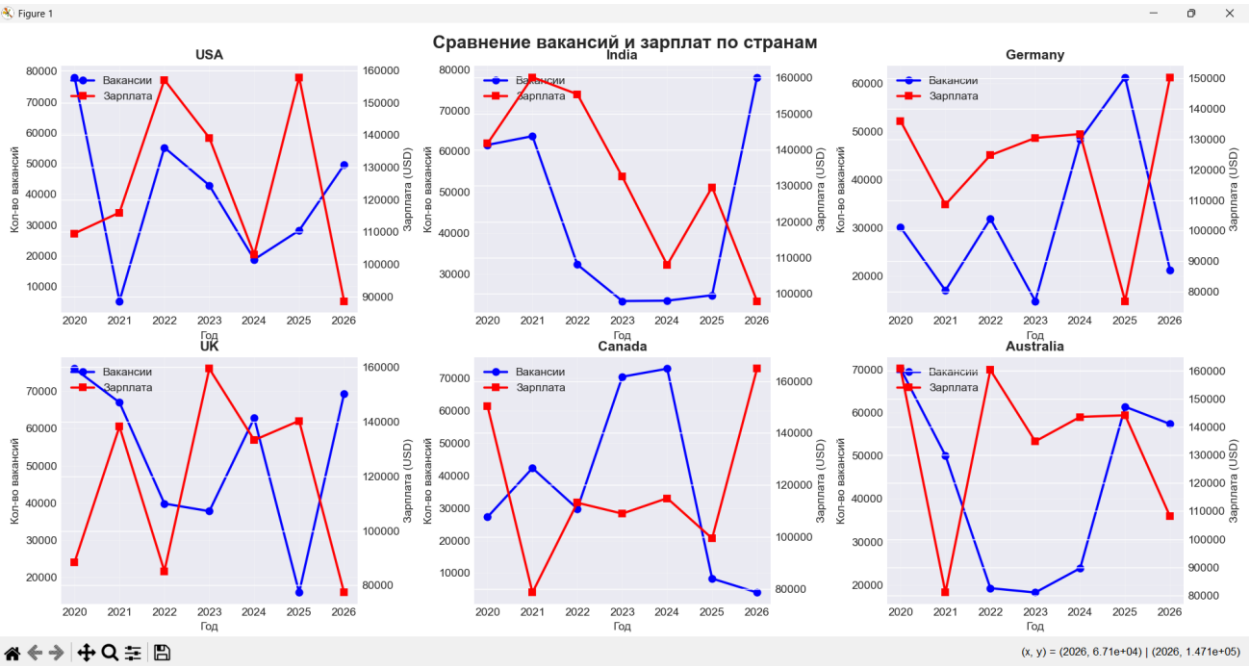


Рис. 6. Третье окно с графиками

Консольный вывод программы предоставляет структурированную аналитическую сводку. Топ-5 записей с самыми высокими зарплатами включают Канаду (2026 г., 164 795 USD), США (2022 г., 157 026 USD и 2025 г., 157 838 USD), Великобританию (2023 г., 159 436 USD) и Индию (2021 г., 160 170 USD). Анализ популярности навыков по всему датасету определяет Python и PyTorch как абсолютных лидеров, за которыми следует GCP. (рис 7).

```
=====
АНАЛИТИЧЕСКАЯ СВОДКА:
=====

Топ-5 самых высоких зарплат:
country year  avg_salary_usd
Canada 2026      164795
Australia 2020    160724
Australia 2022    160260
India 2021      160170
UK 2023        159436

Самые популярные навыки за весь период:
top_skill
PyTorch      7
GCP          5
Python       5
AWS          4
Scikit-learn 4
R            4
NLP          3
Azure        3
SQL          3
TensorFlow   2
Computer Vision 2
Name: count, dtype: int64
```

Рис. 7 Аналитическая сводка

Средние показатели по странам за семилетний период показывают, что Индия лидирует по средней зарплате (135 130 USD), а Великобритания – по среднему количеству вакансий (49 830). Анализ общих трендов по годам выявляет пик средней зарплаты в 2022 году (135 809 USD) и максимальный процент удаленной работы в 2020 году (68%) (рис. 8).

Средние показатели по странам:			
	total_ai_jobs	avg_salary_usd	remote_percentage
country			
Australia	42900.71	133250.71	55.43
Canada	36409.00	118605.86	65.43
Germany	32154.00	122667.71	62.00
India	43901.71	132188.71	45.00
UK	52686.86	117478.57	61.43
USA	39736.57	124475.86	49.29

Общие тренды по годам:			
	total_ai_jobs	avg_salary_usd	remote_percentage
year			
2020	343314	131143.33	67.67
2021	245414	113774.33	48.33
2022	208282	132642.00	58.67
2023	207551	134216.33	48.33
2024	250318	122411.00	50.00
2025	199930	124712.00	65.00
2026	279713	114546.33	57.00

Рис. 8. Средние показатели по странам и общие тренды по годам в сводке

4 Выводы

Практическое применение библиотеки Seaborn подтвердило её высокую эффективность для комплексной визуализации и анализа динамики мирового рынка AI-специалистов. Результаты исследования выявили существенную географическую дифференциацию уровня зарплат с лидирующими позициями Канады, США и Индии в отдельные периоды. Визуализация не обнаружила прямой зависимости между процентом удаленной работы и размером компенсации, что указывает на более сложную природу формирования доходов в данной сфере. Анализ востребованных навыков установил доминирующую роль Python и PyTorch как ключевых технологических компетенций. Разработанная методология демонстрирует потенциал инструментов визуализации данных для мониторинга конъюнктуры рынка труда и может быть адаптирована для анализа других профессиональных отраслей и географических регионов.

Библиографический список

1. Mann M. Global AI and Data Science Job Market (2020-2026): датасет // Kaggle. 2024. Файл country_ai_trends.csv. URL: https://www.kaggle.com/datasets/mann14/global-ai-and-data-science-job-market-20202026/download?datasetVersionNumber=1&select=country_ai_trends.csv
2. Фарайдунов О.К., Амонбекова З.А. Анализ данных в среде языка

- программирования Python для решения экономических задач // Финансово-экономический вестник. 2024. №4-1 (43). С. 91-98. URL: <https://www.elibrary.ru/item.asp?id=80380502>
3. Шарапов Э.С., Корнева Д.С. Статистический анализ датасета заработных плат в сфере Data Science: Факторы влияния и тенденции рынка труда // Научный лидер. 2025. №31 (232). С. 55-56. URL: <https://www.elibrary.ru/item.asp?id=82706687>
 4. Олимов Н.А. Использование искусственного интеллекта в цифровой экономике // Открытое образование. 2025. Т.29. № 2. С. 46-54. URL: <https://www.elibrary.ru/item.asp?id=83174980>
 5. Таран Е.А. Цифровой капитализм и цифровые общественные блага // Экономика и управление инновациями. 2022. №1 (20). С. 13-19. URL: <https://www.elibrary.ru/item.asp?id=48460651>
 6. Sial A. H., Rashdi S. Y. S., Khan A. H. Comparative analysis of data visualization libraries Matplotlib and Seaborn in Python //International Journal. 2021. Т. 10. - №. 1. С. 277-281. URL: <https://www.academia.edu/download/65736020/ijatcse391012021.pdf>
 7. Rajpurkar P., Lungren M. P. The current and future state of AI interpretation of medical images //New England Journal of Medicine. 2023. Т. 388. №. 21. С. 1981-1990. URL: <https://www.nejm.org/doi/full/10.1056/NEJMra2301725>
 8. VanderPlas J. Python Data Science Handbook: Essential Tools for Working with Data. 2nd ed. O'Reilly Media, 2022. URL: <https://www.academia.edu/download/62974298/pythondatasciencehandbook20200415-66956-1noqww2.pdf>
 9. World Economic Forum. The Future of Jobs Report 2023. WEF, 2023. URL: <https://www.voced.edu.au/content/ngv:96765>