DOI: 10.15514/ISPRAS-2025-37(4)-27

Development of Knowledge-based Intelligence for Sustainability Assessment of Russian Regions

[1] *D.S. Fedoseev, ORCID: 0009-0003-0118-6425 <fs159@mail.ru>*
[2] *A.D. Neroslov, ORCID: 0000-0002-8870-1607 <9091069060@mail.ru>*
[1] *V.V. Lanin, ORCID: 0000-0002-0650-2314 <vlanin@hse.ru>*

[1] *HSE University,*
38, Studencheskaia St., Perm, 614070, Russia.

[2] *Perm National Research Polytechnic University,*
29, Komsomolsky Prospekt, Perm, 614990, Russia.

Abstract. The paper presents the development of a Knowledge-based Intelligence for Sustainability Assessment (KISA) system for the comprehensive assessment of the sustainability of Russian regions, which uses a large language model (LLM) with retrieval-augmented generation (RAG) technology and Rosstat data. KISA automatically selects relevant indicators based on users' textual queries, determines their weights, and calculates regional ratings, overcoming the limitations of traditional methods associated with high resource costs, subjectivity, and low adaptability. The system reduces the time required for rating formation to 10 minutes – 140 times faster than existing approaches; financial costs are reduced by a factor of 16 due to the minimization of expert participation. The agreement with expert evaluations is 68%, confirming the validity of the method. KISA provides a web interface with map visualization, enhancing flexibility in analysis; the possibility of improvement through the addition of new sources ensures the continuous incorporating experts' experience. The results of the study contribute to the improvement of regional sustainability assessment and can be used in management decision-making.

Keywords: sustainability assessment; Russian regions; artificial intelligence; large language model LLM; retrieval-augmented generation RAG.

For citation: Fedoseev D.S., Neroslov A.D., Lanin V.V. Development of Knowledge-based Intelligence for Sustainability Assessment of Russian Regions. Trudy ISP RAN/Proc. ISP RAS, vol. 37, issue 4, part 2, 2025, pp. 207-218. DOI: 10.15514/ISPRAS-2025-37(4)-27.

Acknowledgements. This work was supported by the Ministry of Science and Higher Education of the Russian Federation (project No. FSNM-2024-0005).

Разработка интеллектуальной системы на основе знаний для оценки устойчивости российских регионов

[1] *Д.С. Федосеев, ORCID: 0009-0003-0118-6425 <fs159@mail.ru>*
[2] *А.Д. Нерослов, ORCID: 0000-0002-8870-1607 <9091069060@mail.ru>*
[1] *В.В. Ланин, ORCID: 0000-0002-0650-2314 <vlanin@hse.ru>*

[1] *Национальный исследовательский университет «Высшая школа экономики», Россия, 614070, г. Пермь, ул. Студенческая, д. 38.*

[2] *Пермский национальный исследовательский политехнический университет, Россия, 614990, г. Пермь, Комсомольский проспект, д. 29*

Аннотация. В статье представлено описание процесса разработки системы интеллектуальной оценки устойчивости на основе знаний (KISA) для комплексной оценки устойчивости российских регионов, которая использует большую языковую модель (LLM) с технологией дополненной генерации поиска (RAG) и данные Росстата. KISA автоматически отбирает релевантные показатели на основе текстовых запросов пользователей, определяет их веса и рассчитывает региональные рейтинги, преодолевая ограничения традиционных методов, связанные с высокими ресурсными затратами, субъективностью и низкой адаптивностью. Система сокращает время, необходимое для формирования рейтинга, до 10 минут – в 140 раз быстрее существующих подходов; финансовые затраты снижаются в 16 раз за счет минимизации участия экспертов. Согласованность с экспертными оценками составляет 68%, что подтверждает валидность метода. KISA предоставляет веб-интерфейс с картографической визуализацией, повышая гибкость анализа; возможность совершенствования посредством добавления новых источников обеспечивает непрерывное включение экспертного опыта. Результаты исследования способствуют совершенствованию оценки региональной устойчивости и могут использоваться при принятии управленческих решений.

Ключевые слова: оценка устойчивости; российские регионы; искусственный интеллект; большие языковые модели LLM; технология RAG.

Для цитирования: Федосеев Д.С., Нерослов А.Д., Ланин В.В. Разработка интеллектуальной системы на основе знаний для оценки устойчивости российских регионов. Труды ИСП РАН, том 37, вып. 4, часть 2, 2025 г., стр. 207–218 (на английском языке). DOI: 10.15514/ISPRAS–2025–37(4)–27.

Благодарности. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (проект № FSNM-2024-0005).

1. Introduction

The problem of sustainable development of the regions of the Russian Federation has been an urgent interdisciplinary task that requires a comprehensive approach. Regional sustainability is defined as the ability of a region to maintain the quality of life of the population, resist negative impacts, adapt to changes and utilize opportunities for long-term development [1-2].

According to the World Bank [3], Russia ranks third in terms of interregional inequality in Europe and Central Asia, which, in combination with Presidential Decree No. 309 “On the National Development Goals of the Russian Federation for the Period up to 2030 and in Perspective up to 2036” [4], emphasizes the need for a systematic regional sustainability assessment (SA) to reduce the socio-economic gap between regions. This conclusion is confirmed by the studies of Russian scientists [5-6].

At present, assessments are predominantly based on authors’ methods and expert approaches implemented through information systems (IS) for calculating indexes and compiling ratings of regions. Nevertheless, as Ramos notes [7], existing solutions face serious limitations.

The key problems include significant resource costs associated with attracting experts, as well as, the subjectivity and narrow focus of methods that often address a single area, ignoring a comprehensive approach. Moreover, these systems show low adaptability to modern challenges such as climate change and geopolitical crises due to the rigid structure of the indicators used.

The solution to these problems is the development of the Knowledge-based Intelligence for Sustainability Assessment (KISA) system, which utilizes large language model (LLM) with retrieval-augmented generation (RAG) technology and Rosstat data for regional SA. Based on users' text queries, the system automatically selects relevant indicators, determines their weighting coefficients and calculates the integral rating of regions, overcoming the limitations of traditional methods.

2. Problem Statement

The aim of this research is to develop the KISA system capable of generating ratings of Russian regions based on individual user requests, thereby enhancing the objectivity of assessment while reducing the resource intensity of the process.

The project considers only the constitutionally enshrined constituent entities of the Russian Federation (as of 2020), with the analysis based on Rosstat's open data on environmental, economic, and social indicators for the period 2000-2024. While Rosstat provides the most comprehensive official statistics available, it is important to acknowledge potential limitations such as reporting delays and methodological changes, though these constraints remain consistent across all analytical approaches to regional assessment.

Project success criteria:

- Achieving at least an 80% similarity with expert assessments using quality metrics.
- Reducing resource costs for assessment: time by a factor of 7, finances by a factor of 5 compared to traditional expert methods.

The KISA system will not only ensure increased efficiency, adaptability and objectivity of regional SA, but will also simplify the decision-making process for stakeholders, such as government agencies, think tanks and business structures.

3. Related Works

This section focuses on ways to overcome the identified limitations: enhancing the adaptability of evaluation criteria, ensuring objectivity, and optimizing costs compared to traditional approaches. Special attention is paid to the prospects of using artificial intelligence (AI) technologies to address these problems.

3.1 Expert Methods for Sustainability Assessment

Methods for SA usually include sets of statistical indicators and algorithms for calculating an integral index reflecting the level of sustainability of the territory. Numerous such methods have been proposed by both foreign researchers [8-10] and Russian researchers [11-13]. However, many of them have disadvantages: low adaptability due to the use of static indicators and high resource consumption, requiring large research teams.

A methodological review by Lindfors [14] examined the biases that arise when applying such methods, where an excess or deficiency of criteria can distort the evaluation results. This emphasizes the need to develop more efficient, objective and adaptive approaches.

The next subsection examines existing IS for SA, most of which are based on the methods described previously, thus inheriting their limitations.

3.2 Information Systems for Sustainability Assessment

IS for SA were analyzed to identify their strengths and weaknesses for project development.

Three systems were considered: SberAnalytics "Monitoring of regional economies" [15], which provides economic indicators for analyzing regions; Foresight Analytics Platform [16], which has a modular architecture and analytical models; CSA-system [17], which focuses on environmental sustainability of buildings.

These IS applied expert methods for SA, which resulted in the problems of resource intensity, subjectivity and low flexibility. For example, SberAnalytics requires regular participation from experts to customize indicators, and the rigid set of these indicators makes it difficult to adapt to new conditions. Although Foresight allows customization of indicators, the final decisions still depend on expert opinion.

Previously, the current authors (Fedoseev et al.) developed an IS for regional SA [18], which also inherited the limitations of expert methods: the user must independently select indicators and set weighting coefficients, thus preserving the outlined problems.

Despite the advantages of IS in data processing [19], existing solutions only partially solve the problems. The development of an intelligent system is needed to overcome these limitations. The use of AI techniques is promising for creating an effective SA tool [20].

3.3 Application of Artificial Intelligence for Sustainability Assessment

The application of AI in the SA area significantly improves the efficiency of solutions [21]. In particular, the use of Russian-adapted LLMs such as RuAdaptQwen-2.5 [22] with the Chain of Thought approach [23] will allow the KISA system to flexibly customize criteria based on users' textual queries, thereby increasing the adaptability of the system.

The addition of RAG to the LLM will provide the KISA system with the ability to utilize validated data when selecting indicators and setting weights, ensuring a high degree of objectivity in SA. Studies in related fields confirm the effectiveness of this approach: Bronzini et al. [24] applied it to extract information from sustainability reports, and Arslan et al. [25] developed a chatbot for the sustainable energy transition. However, an integrated AI-based regional SA system has not yet been developed.

Systems combining LLM and RAG show significant superiority over traditional methods. According to a study by Ren et al. [26], LLMs outperform humans in all environmental and economic metrics, reducing costs by more than 150 times (from $12.1 to $0.08), which solves the problem of high resource costs associated with expert involvement.

Thus, integrating LLM and RAG into the KISA system will create the first comprehensive system for regional SA, overcoming the key limitations of traditional methods: low adaptability, subjectivity, and high resource costs.

4. System Design

For the systematized storage and processing of statistical data obtained from Rosstat, a database in the form of an ER-diagram was designed (Fig. 1). The central entity “indicator” links numerical indicators with a specific region and time period. The “indicator_type” entity contains the names of specific indicators (e.g., “Unemployment rate”), which are categorized through the “indicator_type_group” entity (e.g., “Economic sphere”). The entity “external_service_data” allows for the integration of information from external sources.

The KISA system is designed using a modular multi-layered architecture with clear separation of responsibilities between components: web interface, controllers, services, and repositories. This approach ensures independent modules and facilitates maintenance and scalability of the application.

The C4 model, which divides the system into Context, Container, Component and Code levels, was used as a design methodology [27].

The component-level architecture is represented in the diagram (Fig. 2), while the detailed description of the main modules and their functions is provided in Table 1.

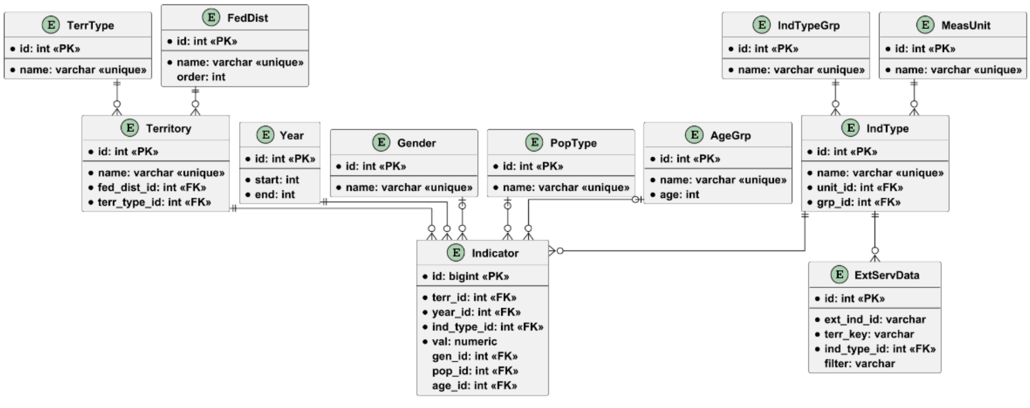

Fig. 1. Entity-Relationship Model of the Database.

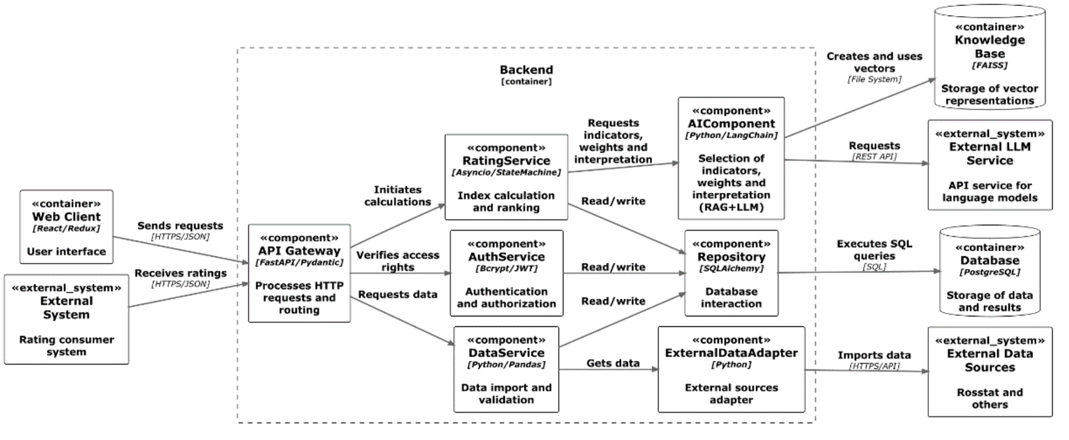

Fig. 2. Backend Components Architecture in C4 Model.

Table 1. Main Modules and Functions.

Module	Functionality
Web Client	User Interface in React
API Gateway	HTTP request processing and routing
AuthService	User authentication and authorization
DataService	Import, validation and structured storage of statistical data using ETL approach
RatingService	Data normalization, removal of outliers, calculation of indexes by weighted sum of indicators and compilation of region rankings
AIComponent	Dynamic selection of relevant indicators and their weights based on text queries using integrated LLM and RAG technologies
Repository	Data access abstraction and database interaction

DataService uses the Extract, Transform and Load (ETL) approach to integrate open statistical data of Rosstat, which ensures a uniform format and correctness of input data [28]. This is important for data unification and subsequent analysis, although possible incompleteness or inaccuracy of the original information should be taken into account.

The algorithm for generating the ratings includes (Fig. 3):

1. User input of a text query defining the SA parameters.
2. Application of RAG technology to extract relevant scientific data from the knowledge base.
3. Transmission of the context, the user's query, and the full list of indicators to the LLM.
4. Automatic selection of relevant indicators by the LLM and determination of their weighting coefficients.
5. Extraction of selected statistical data from PostgreSQL and calculating the integral rating of regions.
6. Visualization of the results with cartographic representation.

The result is a rating reflecting the relative position of each region in terms of regional sustainability.

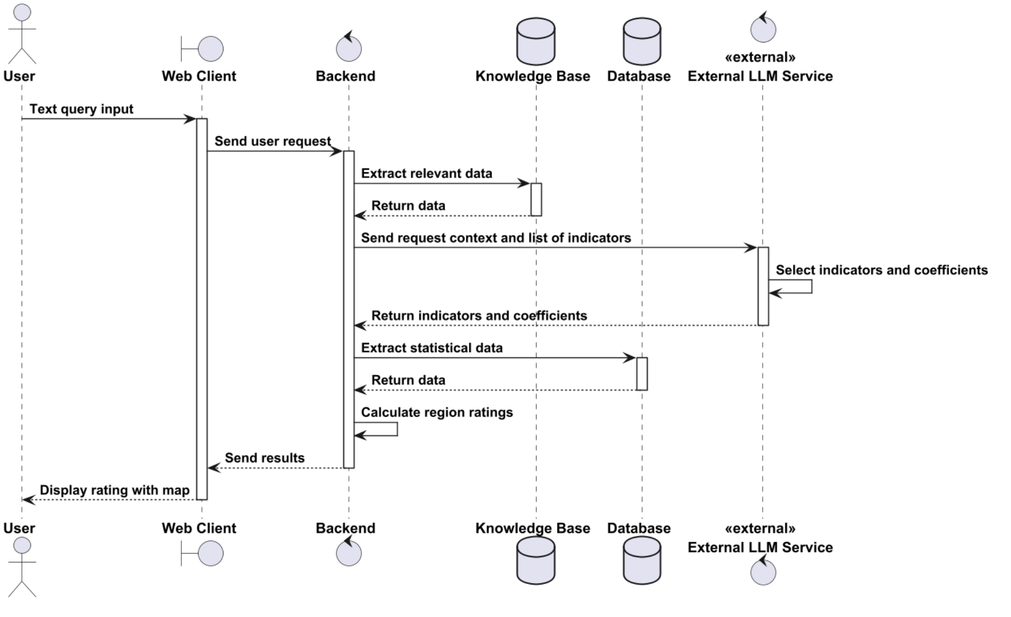

Fig. 3. Rating Formation Process.

5. Implementation

For the development of the KISA system prototype, a technology stack was chosen, including Python, FastAPI for creating a RESTful API [29], and SQLAlchemy for working with a PostgreSQL database. The frontend is developed using React to create an interactive user interface with dynamic graphs and map-based visualization.

LangChain and OpenAI libraries are used to implement RAG and LLM methods. The open-source RuAdaptQwen-2.5 (32 billion parameters), which uses the Chain-of-Thought approach to improve query efficiency, was chosen as the LLM. However, the use of such models requires significant computational power, which may limit query-processing speed. This limitation can be overcome by using powerful servers.

The developed prompts follow effective engineering principles: clear task definition, structured input data presentation, explicit step-by-step reasoning requirements, and formalized output format

(Fig. 4). The system prompt defines the model's role as a data analysis expert and establishes basic operational principles. The user prompt includes a specific task (selecting up to a given number of indicators), user query, knowledge base context, available indicators list, and required response format with examples.

````
## Task:
1. Select up **to {max_inds}** indicators that correspond
to the query.
2. Specify the influence: **1** (positive) or **-1**
(negative).
3. Distribute weights so that the most significant
indicators receive the highest weight, and the least
significant ones receive the lowest.

## User query:
{query}
{context_block}
## List of indicators with ID:
{indicators_text}

## Response format:
1. Explain your choice and distribution of influence and
weights step by step.
2. Provide a JSON response. Example:
```json
{example}
```
````

Fig. 4. Prompt Template Structure for Indicator Selection

For the RAG knowledge base, publications from leading socio-economic journals indexed in Scopus and Web of Science were selected. A total of 243 scientific journals were loaded, covering a wide spectrum of research in regional economics, sustainable development, and methods for assessing territorial entities. The publications were cleaned using regular expressions and segmented into chunks, resulting in 398,290 text fragments. These fragments were then stored in the FAISS vector database using embeddings [30].

KISA integrates RAG technology to access scientific literature for regional sustainability assessments, expanding domain knowledge without costly retraining. By retrieving information from published research, the system bases indicator selection on established methodologies rather than pre-trained knowledge. This integration addresses three critical challenges: resource intensity is reduced through automated analysis, adaptability improves as the LLM flexibly selects indicators based on user text queries, and subjectivity decreases as evaluations rely on published research rather than individual judgment.

6. Evaluation

For the evaluation of the KISA system prototype, an expert commission was formed. The commission consisted of 11 members, including 7 Doctors of Science and 4 Candidates of Science in economics, engineering, history, and social sciences. The commission was tasked with creating three text-query assessment methods based on a set of indicators and the weighted-sum formula used in the KISA system.

The results of the expert assessment and the KISA system were compared. The Recall metric (the ratio of the number of matching indicators to the number of indicators selected by the expert) was used to measure the similarity of the selected indicators, and the MAPE (Mean Absolute Percentage Error) metric, which evaluates the average percentage error between the weights, was used to compare the indicator weights.

The average similarity of the results was 68% (Fig. 5). It is worth noting that a 100% match was not expected, as it is impossible to claim that one method is unequivocally better than the other. The indicators chosen by KISA and experts were similar in meaning but differed in naming. For example, the expert commission chose “Level of education”, while KISA chose “Growth rate of education level”. Although the scope of assessment of the indicators is the same, the approaches to assessment differ, so the results are not completely the same.

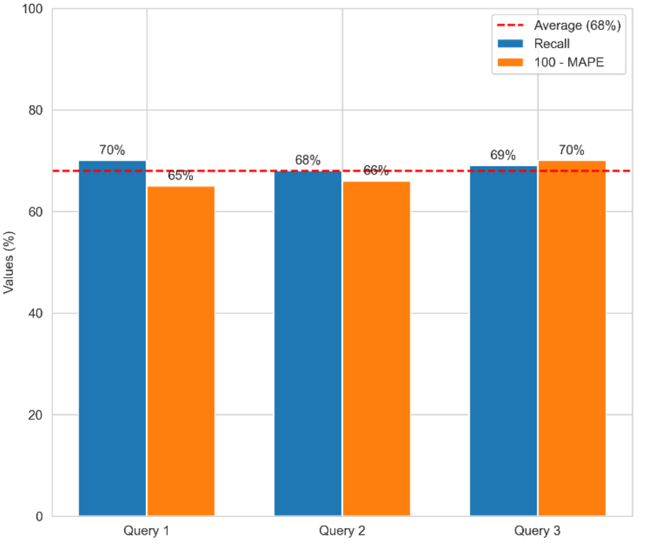

Fig. 5. Comparison of Recall and MAPE across Queries

A straightforward comparison of the system and the experts is complicated, as their assessments may be subjective and reflect different points of view. The similarity of the results also depends on the composition of the expert commission, the specific query, and the content of the knowledge base.

Now, consider comparing the resource costs of the KISA system and the expert commission. According to HeadHunter, an hour of an expert’s work costs 1,000 rubles. An expert commission of 11 people spent three days (24 hours) to create one method, which amounted to 264,000 rubles and required the participation of the entire group. The development of the KISA system took four months with a workload of 160 hours per month. The average salary of a software engineer is 100,000 rubles per month, so the costs totaled about 400,000 rubles.

KISA generates regional ratings on request in 10 minutes, including data input, which saves both financial and time resources. Time costs are reduced by a factor of 140: experts spent 1,440 minutes, and the system completes the same work in 10 minutes. This allows the system to pay for itself after the first request.

Financial costs are reduced 16-fold when using KISA. An expert commission can create a maximum of six methods per month, which costs 1.58 million rubles. At the same time, KISA’s maintenance and infrastructure costs are about 100,000 rubles per month. Thus, the use of the KISA system significantly reduces both time and financial costs, ensuring the efficiency and accessibility of the tool for regional SA.

7. Conclusion

As a result of this study, a working prototype of the KISA system has been developed that addresses three key problems of existing approaches:

1. Resource intensity. Automation of data collection and processing through ETL pipelines and the use of the RuAdaptQwen-2.5 LLM has reduced the time required for rating generation to 10 minutes, which is 140 times faster. Financial costs are reduced 16-fold due to the minimization of expert involvement.
2. Low adaptability. The system interprets text queries in natural language and adapts criteria for specific analysis purposes, providing flexibility and aligning with user needs.
3. Subjectivity. Automated selection of indicators and their weights based on RAG technology eliminates subjective expert assessments. A transparent algorithm based on scientific data increases the reproducibility and validity of the results. The average similarity between the results and expert assessments is 68%.

KISA facilitates the formation of regional ratings by calculating the final regional sustainability index and ranking, as well as visualizing the results through an interactive web interface with a map display.

The key feature of KISA is the possibility of continuous improvement by adding new scientific sources to the knowledge base, allowing the system to be updated and incorporating experts' experience.

The KISA system's database is registered in Rospatent [31]. The computer program has also been registered and received the corresponding certificate [32].

In the future, plans include introducing the function of forecasting sustainability indicators to provide a complete picture of regional development for the next 10 years. Moreover, additional formulas for calculating the sustainability index are expected to be implemented.

The obtained results are important for improving the process of regional SA and can be used to support managerial decisions. Further research will be aimed at expanding the functionality of the system and updating it with new data.

References

[1]. Перфилов В. А. Сущность и типы устойчивости развития региональных социально-экономических систем // ПСЭ. №2., 2012 г., стр 264-266. / Perfilov V. The essence and types of sustainability of regional socio-economic systems. Problems of the modern economy, 2012, vol. 2, pp. 264-266 (In Russian).

[2]. Jovovic R., Draskovic M., Delibasic M., Jovovic M. The concept of sustainable regional development–institutional aspects, policies and prospects. Journal of International Studies, 2017, vol. 10, No. 1, pp. 255-266.

[3]. Inside the World Bank's new inequality indicator: The number of countries with high inequality. [Online]. Available at: https://blogs.worldbank.org/en/opendata/inside-the-world-bank-s-new-inequality-indicator--the-number-of-, accessed: 22 May 2025.

[4]. О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года: Указ Президента РФ от 7 мая 2024 г. № 309 / Decree of the President of the Russian Federation. No. 309. "On the national development goals of the Russian Federation for the period up to 2030 and for the future up to 2036" Available at: https://mvd.consultant.ru/documents/1058493, accessed: 22 May 2025 (In Russian).

[5]. Mareeva S. Socio-economic inequalities in modern Russia and their perception by the population. The Journal of Chinese Sociology, 2020, vol. 7, No.1, 10 p.

[6]. Shatalova O., Kasatkina, E. Socio-economic inequality of regions in the Russian Federation: Measurement issues and long-term evaluation. Economic and Social Changes: Facts, Trends, Forecast, 2022, vol. 15, No. 4, pp. 74-87.

[7]. Ramos T. Sustainability Assessment: Exploring the Frontiers and Paradigms of Indicator Approaches. Sustainability, 2019, vol. 11, No. 3, 824 p.

[8]. Zhong R., Pei F., Yang K., Xia Y., Wang H., & Yan G. Coordinating socio-economic and environmental dimensions to evaluate regional sustainability—towards an integrative framework. Ecological Indicators, 2021, vol. 130, 108085 p.

[9]. D'Adamo I., Falcone P., Imbert E., Morone P. Exploring regional transitions to the bioeconomy using a socio-economic indicator: The case of Italy. Economia Politica, 2022, vol. 39, No. 3, pp. 989-1021.

[10]. Ziegler D., Wolff S., Agu A.-B., Cortiana G., Umair M., Durfort F., Neumann E., Walther G., Kristiansen J., Lienkamp M. How to Measure Sustainability? An Open-Data Approach. Sustainability, 2023, vol. 15, No. 4, 3203 p.

[11]. Швец И. Ю., Швец Ю. Ю., Чиж-Гвязда Э. Оценка устойчивого инновационного развития региона // Вестник Ассоциации вузов туризма и сервиса. том 9. № 1. 2015 г., стр. 14-21. / Shvec I., Shvec Ju., Chizh-Gvjazda Je. Assessment of the sustainable innovative development of the region. Bulletin of the Association of Universities of Tourism and Service, 2015, vol. 9, pp. 14-21 (In Russian).

[12]. Яшина Н. И., Яшин С. Н., Вилейшикова А. А. Методический инструментарий оценки социально-экономической безопасности регионов на основе формирования системы целевых показателей производственной и непроизводственной сфер развития регионов. Вестник Нижегородского университета им. Н.И. Лобачевского. Серия: Социальные науки. Вып. 2, № 62, 2021 г., стр. 45-54. DOI: 10.52452/18115942_2021_2_45. / Jashina N. I., Jashin S. N., Vilejshikova A. A. Methodological tools for assessing the socio-economic security of regions based on the formation of a system of target indicators for industrial and non-industrial areas of regional development. Bulletin of the Nizhny Novgorod Lobachevsky University, 2021, vol. 2, No. 62, pp. 45-54 (In Russian). DOI: 10.52452/18115942_2021_2_45.

[13]. Бородин С. Н. Модель оценки устойчивого развития региона на основе индексного метода // Экономика региона. том 19, № 1, 2023 г., стр. 45-59. DOI 10.17059/ekon.reg.2023-1-4. / Borodin S. A model for assessing the sustainable development of a region based on the index method. Economy of the region. 2023, vol. 19, No. 1, pp. 45-59 (In Russian). DOI 10.17059/ekon.reg.2023-1-4.

[14]. Lindfors A. Assessing sustainability with multi-criteria methods: A methodologically focused literature review. Environmental and Sustainability Indicators, 2021, vol. 12, 100149 p.

[15]. Сбер Аналитика: Мониторинг экономики региона / SberAnalytics "Monitoring of regional economies", Available at: https://sberanalytics.ru/products/gossector/mr, accessed: 22 May 2025 (In Russian).

[16]. Форсайт. Аналитическая платформа / Foresight Analytics Platform, Available at: https://www.fsight.ru/platform/, accessed: 22 May 2025 (In Russian).

[17]. Olawumi T., Chan D. Cloud-based sustainability assessment (CSA) system for automating the sustainability decision-making process of built assets. Expert Systems with Applications, 2022, vol. 188, 116020 p.

[18]. Федосеев Д. С., Нерослов А. Д., Ланин В. В. Разработка информационной системы для оценки устойчивости регионов России. Сборник материалов III студенческой научно-практической конференции им. Л.Л. Любимова. 2025 г., стр. 170-175. / Fedoseev D., Neroslov A., Lanin V. Development of an Information System for Assessing the Sustainability of Russian Regions. Collection of materials of the III student scientific and practical conference named after L.L. Lyubimov at HSE, 2025, pp. 170-175, Available at: https://perm.hse.ru/editorial_publishing/Lyubimov_Conference3, accessed: 22 May 2025 (In Russian).

[19]. Oliveira V., Teixeira D., Rocchi L., Boggia, A. Geographic Information System Applied to Sustainability Assessments: Conceptual Structure and Research Trends. ISPRS International Journal of Geo-Information, 2022, vol. 11, No. 11, 569 p.

[20]. Arfanuzzaman M. Harnessing artificial intelligence and big data for SDGs and prosperous urban future in South Asia. Environmental and sustainability indicators, 2021, vol. 11, 100127 p.

[21]. Greif L., Röckel F., Kimmig A., Ovtcharova J. A systematic review of current AI techniques used in the context of the SDGs. International Journal of Environmental Research, vol. 19, No. 1, 2025.

[22]. Tikhomirov M., Chernyshev, D. Improving Large Language Model Russian adaptation with preliminary vocabulary optimization. Lobachevskii Journal of Mathematics, 2024, vol. 45, No. 7, pp. 3211-3219.

[23]. Kojima T., Gu S., Reid M., Matsuo Y., Iwasawa Y. Large language models are zero-shot reasoners. Advances in neural information processing systems, 2022, vol. 35, pp. 2219-22213.

[24]. Bronzini M., Nicolini C., Lepri B., Passerini A., Staiano J. Glitter or gold? Deriving structured insights from sustainability reports via large language models. EPJ Data Science, 2024, vol. 13, No. 41.

[25]. Arslan M., Mahdjoubi L., Munawar, S. Driving sustainable energy transitions with a multi-source RAG-LLM system. Energy and Buildings, 2024, vol. 324, 114827 p.

[26]. Ren S., Tomlinson B., Black R., Torrance A. Reconciling the contrasting narratives on the environmental impact of large language models. Scientific Reports, 2024, vol. 14, No. 1, 26310 p.

[27]. Vázquez-Ingelmo A., García-Holgado A., García-Peñalvo, F. C4 model in a software engineering subject to ease the comprehension of uml and the software. In 2020 IEEE Global Engineering Education Conference, 2020, pp. 919-924.

[28]. Nwokeji J., Matovu R. A systematic literature review on big data extraction, transformation and loading (etl). In Intelligent Computing: Proceedings of the 2021 Computing Conference, 2021, vol. 2, pp. 308-324.

[29]. Narayanan P. Engineering Machine Learning and Data REST APIs using FastAPI. In Data Engineering for Machine Learning Pipelines: From Python Libraries to ML Pipelines and Cloud Platforms, 2024, pp. 323-359.

[30]. Douze M., Guzhva A., Deng C., Johnson J., Szilvasy G., Mazaré P.-E., Lomeli M., Hosseini L., & Jégou H. The faiss library. arXiv preprint. arXiv:2401.08281, unpublished.

[31]. Паздникова Н. П., Нерослов А. Д., Федосеев Д. С. Комплексная оценка территориальной устойчивости и потенциалов (КОТУП): свидетельство о государственной регистрации базы данных № 2024621915 Российская Федерация; № 2024620239; 2024 г./ Pazdnikova N. P., Neroslov A. D., Fedoseev D. S. Certificate of State registration of the database No. 2024621915 Russian Federation. Comprehensive assessment of territorial stability and potentials (CATSP), 2024.

[32]. Нерослов А. Д., Федосеев Д. С., Паздникова Н. П. Комплексная оценка территориальной устойчивости и потенциалов (КОТУП): свидетельство о государственной регистрации программы для ЭВМ № 2025617985 Российская Федерация. 2025 г. / Neroslov A. D., Fedoseev D. S., Pazdnikova N. P. Certificate of state registration of the computer program No. 2025617985 Russian Federation. Comprehensive Assessment of Territorial Sustainability and Potentials (CATSP), 2025.

Информация об авторах / Information about authors

Данил Сергеевич ФЕДОСЕЕВ – выпускник бакалавриата Национального исследовательского университета «Высшая школа экономики», образовательная программа «Программная инженерия». Сфера научных интересов: обработка естественного языка, большие языковые модели, retrieval-augmented generation.

Danil Sergeevich FEDOSEEV – graduate of the Bachelor’s program in Software Engineering at the National Research University – Higher School of Economics. Research interests: natural language processing, large language models, retrieval-augmented generation.

Алексей Дмитриевич НЕРОСЛОВ – студент первого курса магистратуры Пермского национального исследовательского политехнического университета, образовательная программа «Государственное и муниципальное управление». Сфера научных интересов: экономика, экономика регионов России, социально-экономические показатели деятельности субъектов Российской Федерации.

Alexey Dmitrievich NEROSLOV – first-year master’s student at Perm National Research Polytechnic University in the State and Municipal Administration program. Research interests: economics, regional economics of Russia, socio-economic indicators of the performance of the constituent entities of the Russian Federation.

Вячеслав Владимирович ЛАНИН – старший преподаватель кафедры информационных технологий в бизнесе Национального исследовательского университета «Высшая школа экономики». Сфера научных интересов: языки моделирования, предметно-ориентированное моделирование, языковые инструментарии, CASE-средства, системы имитационного моделирования.

Viacheslav Vladimirovich LANIN – Senior Lecturer of the Department of Information Technologies in Business of the National Research University – Higher School of Economics. Research interests: modeling languages, domain specific modeling, language toolkits, CASE tools, simulation systems.