Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference "Dialogue 2025"

April 23–25, 2025

Generating Encyclopedic Articles Based on a Collection of Scientific Publications

George Lee
Lomonosov Moscow State University
lee.george@mail.ru

Natalia Loukachevitch
Lomonosov Moscow State University
louk_nat@mail.ru

Alexey Khokhlov
Lomonosov Moscow State University
khokhlov@polly.phys.msu.ru

Abstract

Generating texts that demand high factual accuracy and strict formatting, such as encyclopedic articles, presents numerous challenges: how and where to gather relevant information, how to structure it into a coherent and well-formatted text, and how to ensure that the compiled article does not contain factual errors. We propose a solution to these problems for the Russian language by developing a system for generating encyclopedic articles that extracts the most recent and relevant knowledge from scientific publications in the online library eLIBRARY.RU and structures it as a single context for input into a generative model. To evaluate both the impact of the extracted knowledge on the content of the final texts and the overall quality of generation, we considered several prompting strategies, some of which do not use the context found in publications, and compared these approaches using automatic metrics and human expert evaluation. We hope that the created framework will become a reliable reference material for scientists researching new and relevant topics in their field of expertise.

Keywords: Encyclopedic Article Generation, Scientific Publications, eLIBRARY.RU, Retrieval-Augmented Generation, Large Language Models

DOI: 10.28995/2075-7182-2025-23-XX-XX

Порождение энциклопедических статей по коллекции научных публикаций

Ли Г. Ф.
МГУ им. Ломоносова
Москва, Россия
lee.george@mail.ru

Лукашевич Н. В.
МГУ им. Ломоносова
Москва, Россия
louk_nat@mail.ru

Хохлов А. Р.
МГУ им. Ломоносова
Москва, Россия
khokhlov@polly.phys.msu.ru

Аннотация

Порождение столь требовательных по фактологической точности и строгости оформления текстов, как энциклопедические статьи вызывает множество трудностей: как и откуда собирать релевантную информацию, как структурировать ее в связный и грамотно оформленный текст, и как удостовериться, что составленная статья не содержит фактических ошибок. Мы предлагаем решение этих проблем для русского языка, разработав систему генерации энциклопедических статей, извлекающую самые последние и релевантные знания из научных публикаций онлайн-библиотеки eLIBRARY.RU и структурирующую их как единый контекст на вход генеративной модели. В целях оценки как влияния извлеченных знаний на содержание итоговых текстов, так и качества генерации в целом мы рассмотрели несколько стратегий промптирования, часть которых не использует найденный в публикациях контекст, и сравнили эти подходы на автоматических метриках и оценках людей-экспертов. Мы надеемся, что созданный фреймворк сможет стать надежным справочным материалом для ученых при исследовании новых и актуальных тем в области их экспертизы.

Ключевые слова: порождение энциклопедических статей, научные публикации, e.LIBRARY.RU, поисковая дополненная генерация, большие языковые модели

1 Introduction

Encyclopedic articles are a valuable source of information, storing knowledge in a concise but informative form. Currently, a wide variety of online encyclopedias have been created and maintained, the most famous and comprehensive of which is Wikipedia. However, articles for such Internet sources are written and edited by people, which creates difficulties in timely updating them with current knowledge.

A solution to this problem is the creation of systems for automatically compiling encyclopedic articles. Over the past decades, many works (Sauper and Barzilay, 2009; Liu et al., 2018; Fan and Gardent, 2022) have been published describing such frameworks specifically for Wikipedia due to its useful properties:
- Huge size of the knowledge base — a large volume of data for creating training collections;
- Clearly defined section structure and the presence of an official guide for formatting article texts;
- Active community of editors, which allows for a more objective assessment of the system's output quality.

One of the important sources for forming Wikipedia articles is scientific articles. It can take a considerable amount of time from the start of active use of a scientific term in a particular field until a Wikipedia article about that term appears, which may not be sufficiently detailed. This work investigates the problem of generating an encyclopedic article for terms not present in Wikipedia based on a collection of scientific articles. At the first stage of the collection analysis, target terms are identified. Publications and sentences containing the term are then extracted. At the final stage, the extracted sentences are fed into generative models to generate the article. In the study, we evaluated several language models and prompts for generation. This task formulation can be described as the Retrieval-Augmented Generation (RAG) task, which aims to improve the quality of content generation by generative language models based on found relevant information.

2 Related Work

In 2009, C. Sauper and R. Barzilay developed an algorithm for "smart" data extraction from scientific texts annotated with thematic sections (Sauper and Barzilay, 2009). This approach uses the Yahoo search engine as a knowledge base. The extraction of relevant context occurs through queries containing the target term and popular topics in the subject area of the term. Despite impressive results for its time, the "smart" extraction algorithm has significant limitations regarding the training dataset, requiring that the corpus texts be annotated by topic.

In the work (Banerjee and Mitra, 2016), the WikiWrite system is described to automatically generate full-length Wikipedia articles. This approach retrieves information about a target entity from the Web and then uses classifiers trained on the content of similar articles to assign web-retrieved content to relevant sections of the Wikipedia article. A key innovation is the two-step integer linear programming model, which synthesizes disjointed content into coherent, human-like summaries for each section. Although the authors reported that the majority of created articles were retained in the online encyclopedia, they noted that it was difficult for the Wikiwrite to meet the requirements of encyclopedic tone of Wikipedia.

In 2018, the Google Brain team developed a framework for generating preambles for English-language articles in the style of Wikipedia formatting (Liu et al., 2018). Unlike the previous approach, the system extracts context both through Internet queries and through the reference section available for each Wikipedia article. The main innovation in their research was the use of an optimized generative model based on the Transformer architecture (Vaswani, 2017), capable of processing contexts up to 11 thousand tokens long. It is worth noting that the task of generating preambles is significantly simpler compared to the task of generating complete articles, which we pursue in this study.

A. Fan and C. Gardent conducted research on generating Wikipedia-style biographies (Fan and Gardent, 2022), in which a BART-like model (Lewis, 2019) writes complete articles sequentially, section by section. Distinctive features of their system include a citation module that preserves references to the documents used in the extracted context and generates a literature section based on them, as well as a cached generation method that reduces the amount of repetitive information between adjacent sections. The main drawback of this framework is the BART model's input size limitation of 1024 tokens, which significantly reduces the potential amount of information transmitted through the extracted context.

Currently, approaches to generating encyclopedic articles based on the application of large language models (LLMs) are actively being researched. In work (Gao et al., 2024), LLMs are tested in the task of generating reviews dedicated to well-known terms in natural language processing, such as *word2vec*. The testing is conducted in zero-shot (task description without examples), one-shot (based on a single example article), and descriptive prompt formats, where the main sections of the review are specified and explanations are given for each section. To improve the quality of the generated articles, the publications referenced in the corresponding Wikipedia articles can be used. The study shows that the GPT-4 model generates high-quality reviews comparable to expert reviews. However, it should be noted that in this case, reviews are generated for well-known concepts for which there is a large volume of literature.

In the work (Shao et al., 2024), the STORM system is described to help write high-quality Wikipedia articles. Upon receiving a topic for writing an article, the STORM system queries the LLM to find related topics, retrieves articles on related topics from Wikipedia, and extracts tables of contents from them.

In the work (Balepur et al., 2023), an approach is considered to generate descriptive texts on a given topic using the IRP (Imitate, Retrieve, Paraphrase) method. The first component, Imitator, plans the future content of the text. The second component, Retriever, searches for texts based on the formulated query. The last component, Paraphraser, adjusts the extracted text to match the style of the text being created.

In 2024, J. Zhang et al. conducted the research on generating full-length Wikipedia articles on recent real world events (Zhang et al., 2024). The authors developed three RAG frameworks: Retrieve-then-Read for reranking related documents from the Web and simply reading the top ones for generation, Plan-Retrieve-Read (PRR) which uses LLMs' planning capabilities and a multi-stage reranking strategy to outline and generate articles section by section, and a fine-tuned Retrieve-then-Read strategy for Wikipedia generation. The evaluation of generated articles, conducted using LLM-based metrics, showed that hierarchical-based methods (PRR) can produce more comprehensive content, while fine-tuned methods achieve better verifiability. Unlike their study, our system uses a different two-stage reranking technique, experiments with a single RAG framework but multiple prompting strategies and utilizes human expert evaluation since LLMs are known to have certain biases when assessing machine-generated texts (Gao et al., 2024).

The novelty of our paper lies in the pioneering generation of Wikipedia articles for the Russian language and exploring dictionary entry generation in the constrained context of scientific publications.

3 System for Generating Encyclopedic Articles

Our study is based on the following principles:

- *Use of Russian-language scientific publications as an information source.* Presumably, the first information about new terms appears in scientific articles, where the authors describe these concepts and set the context for their use. Extracting information from original publications is a challenging task due to the lack of preliminary filtering and formatting by subject area experts, but it significantly speeds up the knowledge accumulation process by eliminating the need for such experts;
- *Generation using large language models (LLMs).* Training generative models independently not only requires a large annotated dataset but often falls short in quality compared to using the best available LLMs. Moreover, modern models can follow user instructions to adjust their output and work with contexts of enormous size. Therefore, we decided to train only the knowledge extraction system and, instead of training a generative system from scratch, to configure the optimal prompting strategy for a ready-made model.

The developed generation system can be logically divided into separate stages (see Figure 1).

At the first stage, a collection of scientific articles is gathered, and terms for which an encyclopedic article should be written are extracted.

At the second stage, for the target term, sentences from the collected scientific articles that are most relevant to the encyclopedic article are extracted, i.e. sentences that convey the most general information about the term. Two principles are used to select such sentences: automatic sentence classification based on a specially trained classifier, as well as a special list of markers.

At the third stage, the extracted sentences are fed into a generative language model along with a prompt that formulates the task of generating an encyclopedic article. As a result of the final stage, an encyclopedic article is generated for the given term.

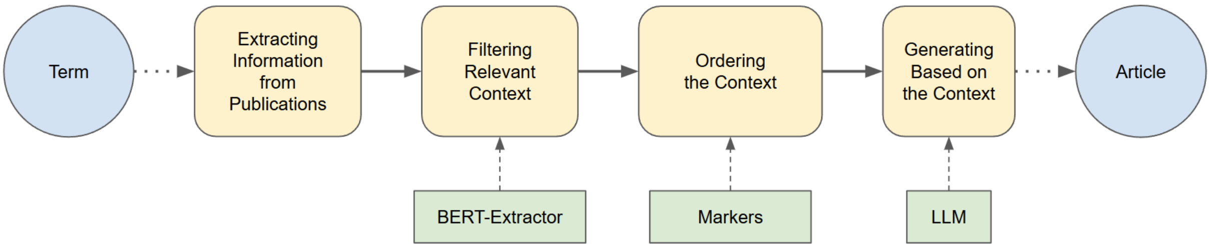


Figure 1: The process of generating an encyclopedic article for a term. Yellow indicates processing stages, green indicates the tools used.

3.1 Sentence Filtering

By a unit of useful information, we mean a separate sentence relevant to the input term. To assess relevance, all candidate sentences from the corpus must first be extracted and processed. The general algorithm for finding sentences can be described as follows:

1. *Selection of articles mentioning the input term*;
2. *Splitting the content of suitable publications into separate sentences*;
3. *Selection of sentences with a given syntactic structure.* For further processing, sentences with a verb predicate and sentences resembling definitions in structure (X is Y) are selected;
4. *Cleaning sentences of insignificant fragments.* Insignificant fragments include section headings, formulas, and hyperlinks, i.e., elements that do not carry significant information but complicate the determination of sentence relevance. Detection of such fragments is done using regular expressions.

The output of the above algorithm is a list of cleaned sentences potentially relevant to the input term. Next, we will consider these stages in more detail.

3.2 Selection of Sentences Relevant for the Encyclopedic Article

Sentences mentioning a term can convey general information about it, useful for inclusion in an encyclopedic article (hereinafter referred to as *encyclopedic sentences*), or specific information related to a particular study. To select encyclopedic sentences, automatic classification based on a transformer neural network model is used. To train the classifier, it was necessary:

1. to create a training dataset based on a given collection of publications;
2. to determine the most suitable model for this dataset and the task.

Training Dataset

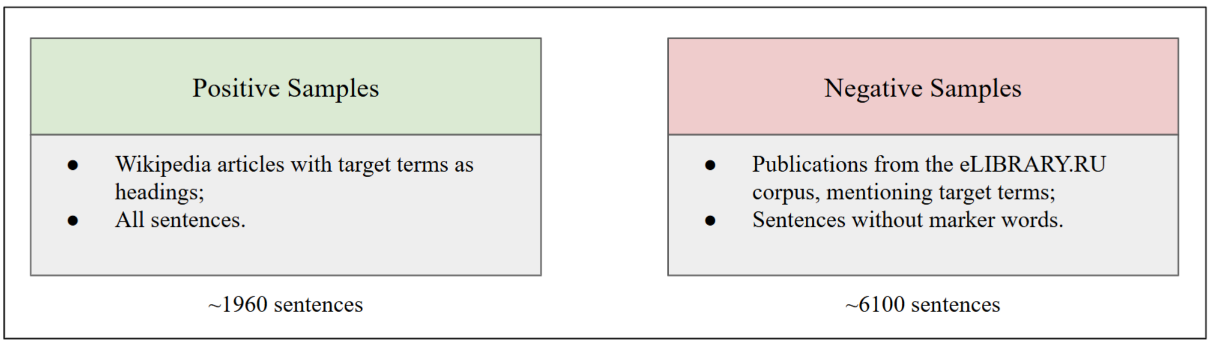


Figure 2: Training data for the task of classifying encyclopedic sentences.

The created dataset consists of automatically annotated sentences for 50 terms that have articles in Wikipedia. Positive sentences are taken from the corresponding Wikipedia articles. The set of negative sentences was formed from scientific articles containing these terms (see Figure 2).

We assumed that the majority of sentences from scientific articles containing the target term relate to a specific study rather than to encyclopedic information. Thus, sentences from articles constitute the set of negative examples for classification. Sentences that potentially could contain encyclopedic information were not included in the negative set. Such sentences usually contain certain markers for conveying basic information about a term and the corresponding concept, i.e. definitin, components and structure, purpose and use, comparisons, as well as markers emphasizing the general significance of the statement contained in the sentence (see Table 1).

Category	**Vocabulary (# words)**	**Representatives**
Definitions	11	вариант, вид, означает, определение, определить, определять, представлять, разновидность, термин, тип, это
Structure	13	включать, включить, входит в состав, компонент, конструкция, модуль, принадлежать, содержат, состав, состоять, структура, часть, элемент
Usage	14	давать возможность, дать возможность, использование, использовать, использоваться, позволить, позволять, предназначить, применение, применить, применять, применяться, служить, удобный инструмент
Distinctive Features	19	больше, больший, меньше, меньший, недостаток, неэффективность, неэффективный, особенность, отличие, отличительная особенность, похож, преимущество, сравнение, сравнивать, сравнительный, сравнить, сходство, эффективность, эффективный
Additional Information	15	В настоящее время, известно, как правило, как принято, общеизвестно, общепринятый, обычно, по большей части, следует считать, считается, таким образом, традиционно, традиционный, часто, чаще всего
Total	**76**	

Table 1: Marker words for encyclopedic sentences, divided by semantic categories.

As a result, a training dataset of 8060 sentences was obtained, in which 1960 sentences are positive examples (filtered sentences from Wikipedia articles), and 6100 sentences are negative examples (filtered sentences from scientific articles). This dataset can be significantly expanded in the future based on the analysis of other terms and corresponding scientific publications and Wikipedia articles.

Comparison of Classifiers for Extracting Encyclopedic Sentences

On the collected training dataset, the quality of several models from the BERT family was evaluated:

- *DeepPavlov/rubert-base-cased* — a model from DeepPavlov based on the multilingual BERT-base, trained on Russian Wikipedia and news data (Kuratov and Arkhipov, 2019);
- *ai-forever/ruSciBERT* — a Russian version of SciBERT (Beltagy et al., 2019), developed by Sber AI in collaboration with the Laboratory of Machine Learning and Semantic Analysis at the AI Institute of Moscow State University (MLSA Lab) for analyzing scientific publications in Russian (Gerasimenko et al., 2022);
- *ai-forever/ruRoberta-large* — a Russian version of the RoBERTa model (Liu, 2019) from the SberDevices team (Zmitrovich et al., 2023);
- *mlsa-iai-msu-lab/sci-rus-tiny* — a small RoBERTa architecture model developed by MLSA Lab for obtaining embeddings of scientific texts in Russian and trained on data from the eLIBRARY.ru

portal;

- *mlsa-iai-msu-lab/sci-rus-tiny3* — an improved and closed version of the previous model.

The characteristics of the classifiers are provided in Appendix A.

The main goal of this stage is to obtain a sentence classifier with minimal error in identifying positive sentences, i.e. sentences carrying encyclopedic information. Within the study, two classification tasks were considered:

1. *Unconditional* — only the sentence itself is fed into the language model, and the model estimates the probability of this sentence to belong to encyclopedic article texts;
2. *Conditional* — as additional information, the target term is fed into the model along with the sentence, separated by the SEP token: ⟨*target term*⟩[**SEP**]⟨*sentence*⟩. In this task, the language model predicts the probability that the sentence belongs to an encyclopedic article dedicated to the target term.

In both cases, contextualized embeddings are generated for the input tokens and the CLS token. For classification, the models *DeepPavlov/rubert-base-cased*, *ai-forever/ruSciBERT*, and *ai-forever/ruRoberta-large* use only the obtained representation of the CLS token, while *mlsa-iai-msu-lab/sci-rus-tiny* and *mlsa-iai-msu-lab/sci-rus-tiny3*, trained to generate semantic embeddings of scientific texts, create an embedding of the entire sentence.

Model	Precision (unconditional)	Precision (conditional)
DeepPavlov/rubert-base-cased	0.9499	**0.9956**
ai-forever/ruSciBERT	0.9618	0.9916
ai-forever/ruRoberta-large	**0.9833**	0.9941
mlsa-iai-msu-lab/sci-rus-tiny	0.9649	0.9642
mlsa-iai-msu-lab/sci-rus-tiny3	0.9338	0.945

Table 2: Precision estimates of various classifiers in the encyclopedic sentences classification task.

It should be noted that the quality of the generative module directly depends on the quality of the extracted sentences, so the precision of classification is a much more significant indicator than its recall. Thus, we can determine the best model using two precision metrics, each corresponding to its classification task (see Table 2).

As can be observed, the *ai-forever/ruRoberta-large* model showed better results than models pretrained on scientific corpora (*ai-forever/ruSciBERT*, *mlsa-iai-msu-lab/sci-rus-tiny*, *mlsa-iai-msu-lab/sci-rus-tiny3*), which can be explained by its size. The final system uses this model because of its high scores on both metrics.

Ordering Encyclopedic Sentences

The extracted encyclopedic sentences can be fed into the generative model without changes. However, we decided to additionally modify the context by ordering the sentences according to the previously mentioned marker words. Presumably, the presence of narrative order in the input context will help the LLM better focus on important fragments for generation.

The sorting of sentences occurs in two stages:

1. *By the semantic segments of the present marker words (primary sorting).* Each semantic segment represents a section of the encyclopedic article. Structurally, it can be described as the following sequence: *definitions, structure, usage, distinctive features, additional information*. As can be seen, the most general categories that can suit terms of different types and different subject areas are chosen as sections. The specified order of sections is used in the primary sorting. If no marker word is encountered in a sentence, it is added to the very end of the sorted sequence;
2. *By the confidence level of the extractor model (secondary sorting).* Almost always after the first stage, segments with many suitable sentences are obtained. To order the context within the sections, a secondary sorting by descending classifier confidence is used.

The relevant sentences ordered in this way will be called the RAG-context of the target term. An example of the joint work of the encyclopedic sentence extraction model and the sorting procedure is provided in Appendix B.

3.3 Context-Based Generation

The responses of the generative RAG-system depend not only on the RAG-context fed into it but also on the used large language model and the type of instructions (prompts) given to the model. The prompt sets the layout of the final text, establishing requirements for content, narrative style, and formatting. The language model uses this layout to generate a response, based on its own and RAG-knowledge, and adhering to the limitations set by its developers. Next, we will consider the LLMs used in the study and the types of prompts.

Used LLMs

We evaluated the quality of generation among three large language models: *Claude-3-sonnet* (Anthropic, 2024), *Meta-Llama-3-70B-Instruct* (Dubey et al., 2024), *Mixtral-8x22b-instruct-v0.1* (Team, 2024). All of them are medium-sized models, providing a balance between speed and quality.

Feature	Claude-3 (Sonnet)	Llama-3 (70B)	Mixtral (8x22b)
Parameter Count	?	70 B	141 B
Context Window (# tokens, 10^3)	200	8.192	64
Knowledge Cutoff	August 2023	December 2023	October 2021
Multimodality	+	-	-
Open Source	-	+	+

Table 3: Features of the used LLMs.

The used LLMs were chosen due to their significantly varying properties (see Table 3), which allowed for a more diversified comparative output and thus simplified the identification of certain patterns.

It is worth mentioning that all the presented models are hosted on different cloud platforms, which may indirectly affect the generated articles. For instance, services for Claude-3 and Mixtral forcibly limit the context window of the models to several thousand tokens, which can lead to unwanted shortening of long articles, while the Llama-3 service does not have such a limitation. Moreover, Llama-3 is the only LLM in our work the workflow of which could be automated through the GradioAPI.

Prompt Instructions Types

Each prompt strategy was created to test some aspect of generation (see Table 4):

- *Basic (Simple prompt)* — general knowledge of the LLM about the target term;
- *Detailed (Detailed prompt)* — model's ability to change the output format according to the user's needs and structure knowledge by sections;
- *Basic/Detailed + RAG* — combining the general knowledge of the LLM with knowledge extracted from scientific articles;
- *Basic + Russian language* — controlling the LLM's proneness to switching to its "default language" when responding to unfamiliar topics;
- *Encyclopedic (Encyclopedic prompt)* — semantic consistency and content completeness of sentences in the RAG-context;
- *Encyclopedic + target term* — model's ability to adapt information from the RAG-context to the target term;
- *Freestyle (Freestyle prompt)* — proficiency in composing a coherent text from the RAG-context.

4 Experiments

4.1 Quality Metrics

The quality of the generated texts can be assessed in different ways. In this work, we used two types of automatic metrics: *absolute* and *relative*. Absolute indicators describe the characteristics of the text, while relative indicators compare the model's prediction with some reference:

Prompt Type	Contents
Basic *+RAG* *+forced lang.*	Твоя задача - ответить на запрос пользователя.[Пиши на русском языке.][Релевантная информация: <RAG-context>] Расскажи о следующем термине: <target term>.
Detailed *+RAG*	Ты - большой эксперт в области физики и химии. Твоя задача - написать основную информацию о запрашиваемом термине. Структурируй ответ по разделам: Определение, Структура, Использование, Отличительные особенности, Дополнительные сведения. Придерживайся указанного порядка разделов. Разделы не являются обязательными. Если считаешь, что на какой-то раздел недостаточно информации, не пиши его. Главный критерий качества твоего ответа - фактологическая точность. Пропускай то, в чем не уверен, не лги пользователю. Старайся писать на русском языке.[При ответе помимо собственных знаний используй предоставленную релевантную информацию. Если считаешь, что какая-то часть релевантной информации не подходит, не используй ее. Релевантная информация: <RAG-context>] Расскажи о следующем термине: <target term>.
Encyclopedic *+target term*	Построй текст энциклопедической статьи на основе заданных предложений.[Статья должна быть посвящена следующему термину: <target term>.] Предложения: <RAG-context> Ответ:
Freestyle	Собери связный текст из заданных предложений. Предложения: <RAG-context> Ответ:

Table 4: Types of prompt instructions and their contents.

- Absolute metrics:
  - *Generation length* — the average number of words in the generated articles;
  - *Latin symbols count* — the average number of Latin characters in the texts, shows how often the LLM switches to its "default language", i.e., English;
- Relative metrics:
  - *ROUGE-Precision* (Lin, 2004) — similarity between two texts based on matching parts.

Matching parts can be:
- ∗ unigrams → *ROUGE-1*;
- ∗ bigrams → *ROUGE-2*;
- ∗ the longest common subsequence → *ROUGE-L* at the level of the entire text, *ROUGE-Lsum* at the level of individual sentences;

Relative metrics were measured when comparing with three variants of reference texts:
1. *The full text of the corresponding Wikipedia article* — similarity between the generated and editor-written articles;
2. *RAG-context* — the activity of using the information extracted from the corpus by the model;
3. *Publications from the eLIBRARY.RU corpus mentioning the target term* — the degree of correspondence of the generation to the latest knowledge in the subject area.

4.2 Input Data Description

The corpus of publications was composed of 15000 scientific articles in physics, downloaded from the electronic library eLIBRARY.RU. During the analysis of the collection, about 4000 articles turned out to be written not in Russian (mainly in English and less often in French) and were removed from the corpus.

From the remaining collection, based on statistical indicators, 20 candidate terms of average frequency were extracted (see Table 5).

Wiki Terms	**Non-Wiki Terms**
1. acoustic emission;	1. pump wave;
2. quantum dot;	2. dynamic recrystallization;
3. internal waves;	3. ionospheric turbulence;
4. Rayleigh wave;	4. quantum beats;
5. gauge bosons;	5. optical rectification;
6. radioactive tracers;	6. parasitic resonance;
7. singlet oxygen;	7. electronic concentration;
8. Stokes number;	8. chaotic synchronization;
9. Rayleigh number;	9. memory function;
10. heavy water.	10. impact layer.

Table 5: Candidate terms for evaluation. Half of them do not have a corresponding Wikipedia article.

4.3 Results

The evaluation was performed for selected terms. The examples of generated articles are provided in Appendix C. The statistics for each variant of the reference text are presented in separate tables, the best indicators for each individual model are underlined, and the best among all three LLMs are highlighted in bold.

Table 6 shows the absolute statistical indicators of the generated articles: the length of the generated text and the number of Latin characters in the texts. From this table, it can be seen that:
- The use of RAG-context and both requirements (to write the text in Russian and for the target term) reduces the size of the final article, regardless of the model and prompt type;
- LLM's occasional switching to the "default language" can be eliminated with the help of RAG-context as well as the explicit requirement to write in Russian;
- On average, Mixtral generates the shortest texts. This is explained by the fact that Claude-3 and Llama-3 usually additionally format their responses (for example, breaking articles into semantic sections and developing each segment into a whole paragraph). Thus, in the case of a detailed prompt, where formatting rules are explicitly set, the difference between the models becomes less noticeable;
- Llama-3 is more prone to switching to English when responding to unfamiliar topics than other models — so, in the case of a basic prompt, it can write the entire article text in English.

Configuration		Generation Length	Latin Count
Claude-3 (Sonnet)	*Basic*	233.85	5.75
	Basic+Lang	213.95	1.15
	Basic+RAG	193.15	8.9
	Detailed	199.9	7.7
	Detailed+RAG	196.6	4.55
	Encyclopedic	239.55	0.5
	Encyclopedic+Term	218.9	0.85
	Freestyle	231.4	0.3
Llama-3 (70B)	*Basic*	246.6	225.35
	Basic+Lang	241.35	20.7
	Basic+RAG	156.6	3.35
	Detailed	**255.85**	16.55
	Detailed+RAG	217.75	1.5
	Encyclopedic	233.8	0.85
	Encyclopedic+Term	226.3	3.9
	Freestyle	249.75	0.7
Mixtral (8x22B)	*Basic*	146.9	0.55
	Basic+Lang	106.15	0.4
	Basic+RAG	138.5	0.45
	Detailed	245.3	1.3
	Detailed+RAG	215.3	1.5
	Encyclopedic	212.4	**0.2**
	Encyclopedic+Term	199.85	0.85
	Freestyle	155.6	0.3

Table 6: Absolute metrics of the generated articles.

Configuration		ROUGE			
		1	2	L	Lsum
Claude-3 (Sonnet)	*Basic*	0.575	0.1786	0.3091	0.5465
	Basic+Lang	0.5773	0.1838	0.3158	0.5441
	Basic+RAG	0.6049	0.2021	0.3307	**0.5771**
	Detailed	0.5808	0.1742	0.3074	0.5533
	Detailed+RAG	0.589	0.1794	0.3109	0.562
	Encyclopedic	0.5351	0.1741	0.2798	0.5003
	Encyclopedic+Term	0.5457	0.1787	0.2933	0.516
	Freestyle	0.534	0.1725	0.289	0.4982
Llama-3 (70B)	*Basic*	0.5537	0.1868	0.3122	0.526
	Basic+Lang	0.523	0.1628	0.2967	0.4981
	Basic+RAG	0.5821	0.2136	0.3496	0.555
	Detailed	0.507	0.1679	0.2867	0.486
	Detailed+RAG	0.509	0.1712	0.2811	0.4828
	Encyclopedic	0.519	0.1639	0.2765	0.4937
	Encyclopedic+Term	0.5315	0.1784	0.2953	0.5052
	Freestyle	0.5144	0.1538	0.2724	0.4819
Mixtral (8x22B)	*Basic*	0.599	0.21	0.3744	0.5673
	Basic+Lang	**0.6055**	0.2008	0.379	0.5741
	Basic+RAG	0.5972	**0.2157**	**0.3876**	0.5652
	Detailed	0.4919	0.1484	0.2765	0.4738
	Detailed+RAG	0.5469	0.1821	0.3136	0.5225
	Encyclopedic	0.5548	0.1947	0.306	0.5226
	Encyclopedic+Term	0.5331	0.1839	0.2961	0.501
	Freestyle	0.5551	0.181	0.3268	0.5219

Table 7: Quality assessment of generation when compared with Wikipedia articles.

Configuration		ROUGE 1	ROUGE 2	ROUGE L	ROUGE Lsum
Claude-3 (Sonnet)	*Basic*	0.3144	0.0649	0.1665	0.2863
	Basic+Lang	0.3756	0.1156	0.2039	0.3427
	Basic+RAG	0.3949	0.104	0.1981	0.3613
	Detailed	0.378	0.1048	0.1986	0.3469
	Detailed+RAG	0.4139	0.1421	0.2161	0.3832
	Encyclopedic	0.524	0.3393	0.314	0.4994
	Encyclopedic+Term	0.6003	0.4443	0.3552	0.5816
	Freestyle	0.6932	0.6529	0.4334	0.6877
Llama-3 (70B)	*Basic*	0.2943	0.0592	0.1638	0.2681
	Basic+Lang	0.3104	0.0653	0.1782	0.2833
	Basic+RAG	0.6277	0.3868	0.4195	0.5958
	Detailed	0.2974	0.0637	0.1621	0.2697
	Detailed+RAG	0.5119	0.314	0.31	0.4829
	Encyclopedic	0.6755	0.5593	0.5103	0.6593
	Encyclopedic+Term	0.5538	0.3834	0.3821	0.5353
	Freestyle	0.9429	**0.935**	**0.8159**	0.9428
Mixtral (8x22B)	*Basic*	0.3847	0.0787	0.2283	0.351
	Basic+Lang	0.4258	0.086	0.2585	0.3915
	Basic+RAG	0.5493	0.3097	0.3716	0.5137
	Detailed	0.2994	0.0597	0.1708	0.274
	Detailed+RAG	0.4518	0.2345	0.2753	0.4229
	Encyclopedic	0.6802	0.5516	0.485	0.6634
	Encyclopedic+Term	0.6463	0.4748	0.4593	0.623
	Freestyle	**0.9653**	0.9074	0.6459	**0.9595**

Table 8: Quality assessment of generation when compared with RAG-contexts.

Configuration		ROUGE 1	ROUGE 2	ROUGE L	ROUGE Lsum
Claude-3 (Sonnet)	*Basic*	0.9492	0.6162	0.7783	0.943
	Basic+Lang	0.9594	0.6588	0.7977	0.9527
	Basic+RAG	0.973	0.6795	0.8261	0.9676
	Detailed	0.9682	0.6569	0.8157	0.9638
	Detailed+RAG	0.9744	0.6615	0.8177	0.9712
	Encyclopedic	0.9911	0.7821	0.8243	0.9869
	Encyclopedic+Term	0.9957	0.8265	0.8545	0.9926
	Freestyle	0.9868	0.8811	0.8496	0.9852
Llama-3 (70B)	*Basic*	0.8654	0.517	0.7227	0.8584
	Basic+Lang	0.9485	0.5788	0.7909	0.9422
	Basic+RAG	0.9831	0.7864	0.8632	0.9801
	Detailed	0.9488	0.5466	0.7648	0.9429
	Detailed+RAG	0.9847	0.717	0.8248	0.9821
	Encyclopedic	0.9936	0.8639	0.8438	0.9922
	Encyclopedic+Term	0.9855	0.7777	0.8334	0.9828
	Freestyle	0.9808	0.9534	0.847	0.9784
Mixtral (8x22B)	*Basic*	0.9658	0.6448	0.8518	0.9621
	Basic+Lang	0.9758	0.6552	0.8924	0.9689
	Basic+RAG	0.992	0.7653	0.8814	0.9886
	Detailed	0.9701	0.6274	0.8163	0.9675
	Detailed+RAG	0.979	0.6989	0.835	0.9761
	Encyclopedic	0.9929	0.8378	0.8558	0.9909
	Encyclopedic+Term	0.9911	0.8309	0.8658	0.9885
	Freestyle	**0.9986**	**0.9693**	**0.9075**	**0.9983**

Table 9: Quality assessment of generation when compared with eLIBRARY.RU publications.

Table 7 shows the similarity measures of the generated texts with Wikipedia articles for terms that have such articles. Table 8 contains the similarity measures of the generated articles with the context extracted from scientific articles. Table 9 shows the similarity measures of the generated articles with the full scientific publications in which the target term was found.

The study of the statistical characteristics of the generated articles revealed the following patterns and properties of LLMs:

- One of the reasons for the high ROUGE scores of Mixtral is the brevity of the texts it generates. This is a known problem of metrics based solely on N-gram precision;
- The use of RAG-context increases the values of ROUGE regardless of the model, slightly when evaluated on Wikipedia articles and significantly on the original publications and the context itself. It is worth noting that adding the requirement to generate for the target term in the encyclopedic prompt improves the metrics only for Claude-3 — this suggests that statistical metrics do not necessarily have to correlate with the level of narrative coherence;
- The encyclopedic and freestyle prompts have the highest scores for ROUGE when evaluated on the RAG-context and original documents, the latter prompting strategy is especially good for these metrics. Such a result is expected, since with this strategy the model essentially compiles a story from the sentences of the RAG-context, simply reordering them and almost not adding new tokens;
- The detailed prompt performed worse than the basic one when evaluated on Wikipedia articles — perhaps the standard generation templates in LLMs are sufficiently similar to the Wikipedia format;
- Claude-3 is the most conservative model, it noticeably uses the proposed context less when generating a response. Perhaps this is influenced by its "closedness" (recall that Llama-3 and Mixtral are open-source models);
- Relatively low ROUGE-2 and ROUGE-L scores when evaluated on Wikipedia articles show that the similarity between the generated and reference articles is observed at most at the level of individual words or short common word sequences. This means that LLMs do not explicitly try to copy Wikipedia when responding to queries.

4.4 Human Evaluation

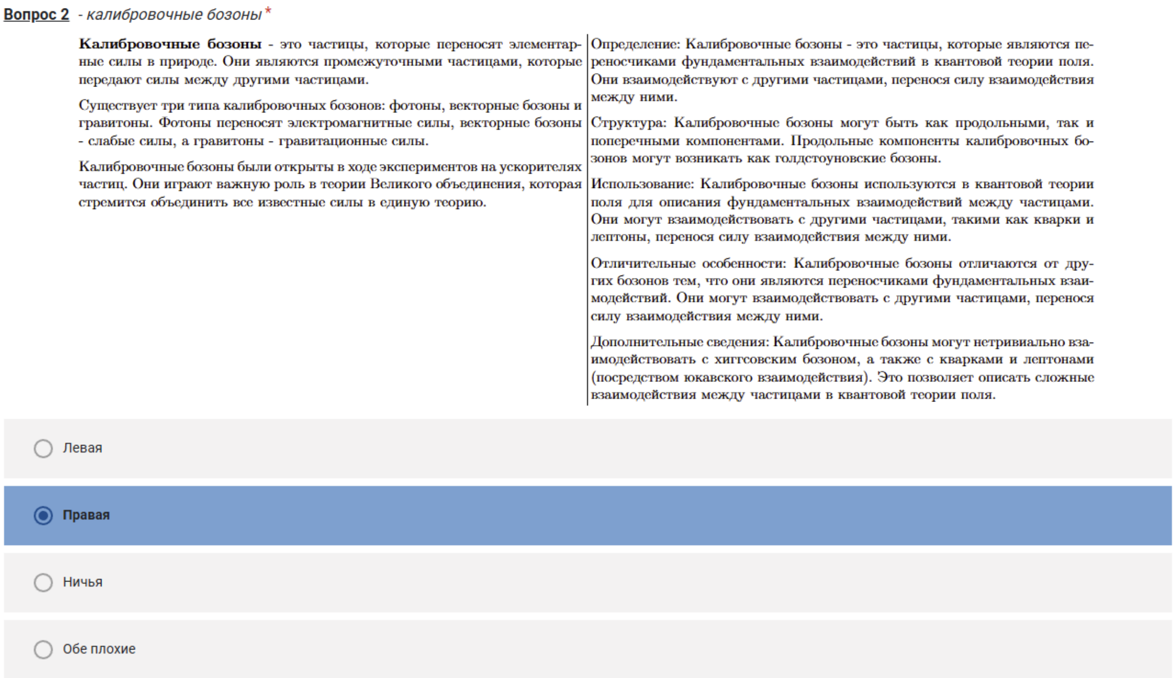
The main drawback of all used automatic metrics is that they evaluate only the statistical characteristics of texts, which say almost nothing about the factual accuracy and logical coherence of what is written. A solution to this problem is the involvement of human experts into the evaluation (Van Der Lee et al., 2019).

As part of our study, we conducted a survey consisting of the following parts:

1. **parallel comparison (60 questions)** — the expert is offered two articles written for the same target term by different models with different prompting strategies (see Figure 3). The task of the respondent is to choose the text that they liked more. There are no specific quality criteria in this task, i.e., the choice is made solely based on the subjective considerations of the expert;
2. **standalone evaluation (40 questions)** — in each question, the respondent is offered one article (see Figure 4)). The task of the expert is to evaluate it according to three criteria:
    - *factual accuracy* — the correspondence of the article's content to real facts;
    - *semantic consistency* — the coherence and logic of the narrative, the completeness of thought;
    - *compliance with the format of an encyclopedic article* — the degree of similarity (in structure and visual design) of the generated article to real encyclopedic articles written by people.

The experts from the Faculty of Physics at Lomonosov Moscow State University were invited to assess the generated articles: 13 experts for the parallel part and 11 for the standalone one. Based on the assessments obtained during the parallel comparison stage, ratings of the best models and prompting strategies were calculated (see Table 10). The scores of the candidates were calculated according to the following procedure:

- *Respondent's choice of candidate i* — the model and prompt corresponding to it receive 1 point. The opponent j does not receive a win;
- *Tie* — both candidates do not receive points;
- *Both are bad* — both opponents are penalized, receiving -1 point.



Вопрос 2 - *калибровочные бозоны* *

Калибровочные бозоны - это частицы, которые переносят элементарные силы в природе. Они являются промежуточными частицами, которые передают силы между другими частицами.

Существует три типа калибровочных бозонов: фотоны, векторные бозоны и гравитоны. Фотоны переносят электромагнитные силы, векторные бозоны - слабые силы, а гравитоны - гравитационные силы.

Калибровочные бозоны были открыты в ходе экспериментов на ускорителях частиц. Они играют важную роль в теории Великого объединения, которая стремится объединить все известные силы в единую теорию.

Определение: Калибровочные бозоны - это частицы, которые являются переносчиками фундаментальных взаимодействий в квантовой теории поля. Они взаимодействуют с другими частицами, перенося силу взаимодействия между ними.

Структура: Калибровочные бозоны могут быть как продольными, так и поперечными компонентами. Продольные компоненты калибровочных бозонов могут возникать как голдстоуновские бозоны.

Использование: Калибровочные бозоны используются в квантовой теории поля для описания фундаментальных взаимодействий между частицами. Они могут взаимодействовать с другими частицами, такими как кварки и лептоны, перенося силу взаимодействия между ними.

Отличительные особенности: Калибровочные бозоны отличаются от других бозонов тем, что они являются переносчиками фундаментальных взаимодействий. Они могут взаимодействовать с другими частицами, перенося силу взаимодействия между ними.

Дополнительные сведения: Калибровочные бозоны могут нетривиально взаимодействовать с хиггсовским бозоном, а также с кварками и лептонами (посредством юкавского взаимодействия). Это позволяет описать сложные взаимодействия между частицами в квантовой теории поля.

- Левая
- **Правая**
- Ничья
- Обе плохие

Figure 3: Example of parallel comparison for the term "calibration bosons".

Вопрос 1 - *квантовое биение*

В последние два десятилетия интерес к квантовым биениям резко возрос в связи с наблюдением биений в квантово-размерных полупроводниковых наноструктурах, рассматриваемых сейчас как наиболее перспективные материалы для электронной и вычислительной техники. Это позволяет более отчетливо регистрировать квантовые биения. Простейшим примером, демонстрирующим квантовые биения в радикальной паре, может служить пара радикалов, не взаимодействующих друг с другом и отличающихся лишь значениями изотропных g-факторов. Они показаны для демонстрации квантовых биений волновой функции. В литературе, однако, описаны лишь единичные примеры, когда при лазерном облучении удавалось наблюдать квантовые биения в рекомбинации пар. Эти динамические переходы (квантовые биения) аналогичны квантовым биениям в других объектах, что позволяет рассматривать спиновую химию в одном ряду с другой новой научной областью - фемтохимией и с некоторыми спектроскопическими методами исследования, в которых также проявляется квантовая интерференция. И в спиновой химии, и в фемтохимии квантовые биения в заселенностях состояний реагентов модулируют выход продуктов реакции.

Фактологическая точность *

Смысловая согласованность *

Соответствие формату энциклопедической статьи *

Figure 4: Example of standalone evaluation for the term "quantum beats".

Looking at the Table 10, it becomes clear that the experts significantly preferred the responses of the "closed" Claude-3 (Sonnet) over the responses of open-source models. As for the prompting strategies, the most successful was the detailed prompt. The use of RAG-context generally makes the generated articles more attractive, but does not guarantee this. Surprisingly, encyclopedic prompts received mostly mixed or even negative feedback. It can be assumed that this is due to both insufficient precision in the formulation of these prompts and the lack of explicitly defined text formatting rules in them.

To determine the optimal generation configuration, the characteristics of the top-2 models and top-3 prompts in the ranking were investigated during the standalone evaluation stage (see Table 11).

From the obtained assessments, the following conclusions can be drawn:

- Claude-3 generates more plausible texts, while Llama-3 generates texts that are more similar in style to encyclopedic articles;
- RAG-context improves the accuracy of the model's responses but worsens their logical coherence. This observation can be explained by the semantic heterogeneity of the extracted relevant sentences, which complicates their integration into a single text.

Rating	Model	Score
1	**Claude-3 (Sonnet)**	**91**
2	Llama-3 (70B)	23
3	Mixtral (8x22b)	-44

Rating	Prompt Type	Score
1	**Detailed**	**50**
2	Basic+RAG	46
3	Detailed+RAG	27
4	Basic	16
5	Basic+Lang	14
6	Encyclopedic+Term	0
7	Encyclopedic	-15
8	Freestyle	-68

Table 10: Models and prompts ratings based on expert evaluation.

Configuration		Factual Accuracy	Semantic Consistency	Format Compliance
Claude-3 (Sonnet)	*Basic+RAG*	4.09	3.45	3.14
	Detailed	4.11	**<u>4.26</u>**	<u>3.82</u>
	Detailed+RAG	**<u>4.36</u>**	4	3.63
Llama-3 (70B)	*Basic+RAG*	3.09	2.91	2.82
	Detailed	3.91	<u>3.82</u>	**<u>3.98</u>**
	Detailed+RAG	<u>3.97</u>	3.51	3.69

Table 11: Results of the standalone evaluation for the best rated configurations.

Thus, the choice of the optimal configuration depends on the importance of each criterion in the task at hand: for the best generation accuracy, Claude-3 with a detailed prompt and RAG-context is suitable; for generating articles that are most similar to encyclopedic ones, Llama-3 with a detailed prompt is recommended; and as a compromise, Claude-3 with a detailed prompt can be chosen.

5 Conclusion

In this work, we developed a system for generating Russian-language encyclopedic articles. The framework is based on the extraction and processing of knowledge from scientific publications of the electronic library eLIBRARY.RU to form a current and relevant context for generative models.

The resulting system can be improved in many ways. Thus, in the future, it is planned to improve the BERT-extractor by including meta-information about the target term in the context (for example, the presence of marker words in the sentence) and replacing the extraction method itself with the "small-to-big" retrieval approach. Both proposals will add additional information for the language model necessary for a more thoughtful and high-quality selection of the RAG-context. We also see the possibility of improving the framework by modifying the optimal prompt strategies — at the moment, the LLM is too inconsistent in its writing style, which is undesirable for scientific articles with fairly strict formatting rules. By adding more specific text formatting rules and several examples of correctly written articles to the prompt, the model will be able to better adjust the generation to real publications.

Acknowledgments

The study was carried out with the financial support of the Interdisciplinary Scientific and Educational School of Lomonosov Moscow State University "Mathematical methods of analysis of complex systems", agreement No. 23-Ш05-21. We thank eLibrary for providing the collections of scientific papers.

References

AI Anthropic. 2024. The claude 3 model family: Opus, sonnet, haiku. *Claude-3 Model Card*, 1.

Nishant Balepur, Jie Huang, and Kevin Chen-Chuan Chang. 2023. Expository text generation: Imitate, retrieve, paraphrase. *arXiv preprint arXiv:2305.03276*.

Siddhartha Banerjee and Prasenjit Mitra. 2016. Wikiwrite: Generating wikipedia articles automatically. // *IJCAI*, P 2740–2746.

Iz Beltagy, Kyle Lo, and Arman Cohan. 2019. Scibert: A pretrained language model for scientific text. *arXiv preprint arXiv:1903.10676*.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Angela Fan and Claire Gardent. 2022. Generating full length wikipedia biographies: The impact of gender bias on the retrieval-based generation of women biographies. *arXiv preprint arXiv:2204.05879*.

Fan Gao, Hang Jiang, Rui Yang, Qingcheng Zeng, Jinghui Lu, Moritz Blum, Tianwei She, Yuang Jiang, and Irene Li. 2024. Evaluating large language models on wikipedia-style survey generation. // *Findings of the Association for Computational Linguistics ACL 2024*, P 5405–5418.

Nikolai Aleksandrovich Gerasimenko, Aleksander Sergeevich Chernyavsky, and MA Nikiforova. 2022. Ruscibert: A transformer language model for obtaining semantic embeddings of scientific texts in russian. // *Doklady Mathematics*, volume 106, P 95–96. Springer.

Yuri Kuratov and Mikhail Arkhipov. 2019. Adaptation of deep bidirectional multilingual transformers for russian language. *arXiv preprint arXiv:1905.07213*.

Mike Lewis. 2019. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.

Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. // *Text summarization branches out*, P 74–81.

Peter J Liu, Mohammad Saleh, Etienne Pot, Ben Goodrich, Ryan Sepassi, Lukasz Kaiser, and Noam Shazeer. 2018. Generating wikipedia by summarizing long sequences. *arXiv preprint arXiv:1801.10198*.

Yinhan Liu. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 364.

Christina Sauper and Regina Barzilay. 2009. Automatically generating wikipedia articles: A structure-aware approach. // *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, P 208–216.

Yijia Shao, Yucheng Jiang, Theodore A Kanell, Peter Xu, Omar Khattab, and Monica S Lam. 2024. Assisting in writing wikipedia-like articles from scratch with large language models. *arXiv preprint arXiv:2402.14207*.

Mistral AI Team. 2024. Cheaper, better, faster, stronger. `https://mistral.ai/news/mixtral-8x22b/`.

Chris Van Der Lee, Albert Gatt, Emiel Van Miltenburg, Sander Wubben, and Emiel Krahmer. 2019. Best practices for the human evaluation of automatically generated text. // *Proceedings of the 12th International Conference on Natural Language Generation*, P 355–368.

A Vaswani. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*.

Jiebin Zhang, Eugene J Yu, Qinyu Chen, Chenhao Xiong, Dawei Zhu, Han Qian, Mingbo Song, Weimin Xiong, Xiaoguang Li, Qun Liu, et al. 2024. Wikigenbench: Exploring full-length wikipedia generation under real-world scenario. *arXiv preprint arXiv:2402.18264*.

Dmitry Zmitrovich, Alexander Abramov, Andrey Kalmykov, Maria Tikhonova, Ekaterina Taktasheva, Danil Astafurov, Mark Baushenko, Artem Snegirev, Vitalii Kadulin, Sergey Markov, et al. 2023. A family of pretrained transformer language models for russian. *arXiv preprint arXiv:2309.10931*.

Appendix A. Properties of Classifiers and Their Training.

Model	Parameter Count	Batch Size	# Epochs	Unconditional Train Time	Conditional Train Time
DeepPavlov/rubert-base-cased	180 M	**16**	3	2.15 min	2.2 min
ai-forever/ruSciBERT	123 M	8	3	2.85 min	2.92 min
ai-forever/ruRoberta-large	**355 M**	4	3	**14.58 min**	**27.26 min**
mlsa-iai-msu-lab/sci-rus-tiny	23 M	**16**	**10**	1.41 min	1.46 min
mlsa-iai-msu-lab/sci-rus-tiny3	23 M	**16**	**10**	1.46 min	1.53 min

Table 12: Characteristics of extractor models and their training process.

Appendix B. Context Filtering by the Extractor Model.

Segment	Context
Definitions	1. Это оборудование позволяет получать большое количество данных в ходе мониторинга и затем проводить полный анализ сигналов акустической эмиссии, чтобы определить изменение показателя b для всего периода исследований. 2. Акустическая эмиссия (АЭ) - явление, сопутствующее многим физическим процессам в твердом теле. 3. Каждый раз при возникновении акустической эмиссии это событие может привести либо к созданию новой элементарной трещины, либо к распространению существующей трещины на некоторую фиксированную величину (квант разрушения).
Structure	4. Поэтому можно считать, что спектральный состав сигнала акустической эмиссии является структурно чувствительной характеристикой.
Usage	5. Применение этого метода ограничивается возможностями выделения сигналов акустической эмиссии из помех, то есть данный метод чувствителен только на заключительных этапах жизни образца, когда размеры дефектов уже достаточно велики и приближаются к критическим значениям.
Additional Information	6. Акустическая эмиссия обычно проявляется при пластической деформации твердых материалов, но в последнее время появились публикации об использовании этого метода и при изучении физико-химических процессов, проходящих в жидкости.
Unidentified	7. Пассивный метод контроля состояния объекта состоит в использовании сигнала акустической эмиссии, излучаемого при деформации объекта контроля, или сигнала естественного акустического фона окружающей среды, отраженного объектом. 8. Полагаем, что такое деление условно и непрерывная и импульсная акустическая эмиссия представляют поток сигналов, различающиеся амплитудой и частотой появления. 9. В технике широко используется спектральный анализ акустической эмиссии шумов, возникающих при эксплуатации агрегатов и устройств, и эмиссии, возникающей при нагружении материала. 10. Метод акустической эмиссии (АЭ) для исследования свойств материалов получил широкое распространение во второй половине прошлого столетия. 11. Таким образом, трещина выступает в качестве рода, в то время как наименьшее событие акустической эмиссии играет роль вида. 12. Такой подход позволяет эффективно сопоставлять эти параметры, особенно при сравнении структуры алюминия, деформации и акустической эмиссии. 13. Разработана методика определения активационных параметров - энергии активации и активационного объема, основанная на анализе акустической эмиссии, зависимой от температуры или механического напряжения.

Table 13: Selected relevant sentences for the term "acoustic emission".

Appendix C. Generation Samples.

WIKI

Калибровочные бозоны – это частицы, переносящие фундаментальные взаимодействия в Стандартной модели физики элементарных частиц. Они являются квантами соответствующих калибровочных полей, описывающих эти взаимодействия.

К калибровочным бозонам относятся:

1. Фотон – переносчик электромагнитного взаимодействия.
2. W^+, W^-, Z^0 бозоны – переносчики слабого ядерного взаимодействия.
3. 8 глюонов – переносчики сильного ядерного (цветового) взаимодействия между кварками.
4. Предполагаемый гравитон – гипотетическая частица-переносчик гравитационного взаимодействия (пока не обнаружен экспериментально).

Основные свойства калибровочных бозонов:

- Являются безмассовыми (кроме W^+, W^-, Z^0) векторными частицами с целочисленным спином.
- Описываются калибровочными теориями, основанными на принципе калибровочной инвариантности.
- Взаимодействуют только с частицами, несущими соответствующий заряд (электрический, цветовой и т.д.)
- Играют фундаментальную роль, будучи переносчиками всех известных фундаментальных взаимодействий (кроме гравитации).

Существование калибровочных бозонов является одним из столпов Стандартной модели и подтверждено многочисленными экспериментами в физике высоких энергий.

NON-WIKI

Хаотическая синхронизация (chaotic synchronization) – это явление подстройки хаотических колебаний в связанных нелинейных динамических системах, при котором их траектории на аттракторах становятся идентичными или функционально связанными.

Существуют несколько основных типов хаотической синхронизации:

1. Полная хаотическая синхронизация – идентичность хаотических траекторий связанных систем в пределах аттрактора.
2. Фазовая хаотическая синхронизация – совпадение только фаз хаотических колебаний, а амплитуды могут различаться.
3. Обобщенная хаотическая синхронизация – функциональная взаимосвязь между хаотическими траекториями систем, но не их идентичность.
4. Частотная хаотическая синхронизация – совпадение лишь основных частот хаотических колебаний.

Хаотическая синхронизация возникает при достаточно сильной диссипативной связи между хаотическими системами. Она может быть односторонней (подавляющая и подавляемая системы) или взаимной.

Явление хаотической синхронизации имеет большое значение для различных приложений, таких как системы передачи информации с хаотической маскировкой, генераторы хаотических сигналов, хаотическая обработка сигналов, нелинейная локация и томография сред.

Хаотическая синхронизация демонстрирует возможность упорядочения хаотической динамики в связанных нелинейных системах и открывает новые пути использования хаотических режимов в технических устройствах.

Figure 5: Generated texts for terms with and without corresponding Wikipedia articles. Claude-3 model with Basic+RAG prompt was used as a configuration.