USING KNOWLEDGE GRAPH IN ADAPTING LAN-GUAGE MODEL ON MATHEMATICAL TEXTS

Anonymous authors

000

001

002 003 004

005

006 007 008

010 011

012

013

014

015

016

018

019 020 021

022

Paper under double-blind review

ABSTRACT

The subject of the study is the problem of adapting language models to scientific subject areas. The issues of expanding language models to mathematical subject areas are considered. It is proposed to use the knowledge graph of the subject area as a tool for 'tuning' the language model. To build the knowledge graph, the ontology of the subject area of the semantic library of mathematical subject areas and their applications LibMeta is used. Navigation through the subject area is carried out using the knowledge graph and is limited by the terminology of the thesaurus and ontology links. This approach allows using the knowledge graph to create a digital assistant in a recommender system, an agent for a language model, and to feed mathematical text data to a language model.

1 INTRODUCTION

Modern large language models (LLM) are being implemented in various information systems to provide a dialogue with the user in natural language. This radically changes the way information is extracted, accelerates the acquisition of necessary information and knowledge in certain cases. The influence of ChatGPT and LLM is enormous. Meanwhile, the university community is discussing the problems of devaluation of scientific work and plagiarism, in connection with the use of knowledge based on ChatGPT Ramil Malinka & all (2023) in the process of obtaining education. The study in this paper is devoted to the acquisition of knowledge in a specific mathematical subject area using LLM.

A well-known problem with using LLM is the difficulty of explaining and veracity of the conclusions, since
 the answer does not indicate the sources on the basis of which the answer is formed. As a rule, such sources
 cannot be specified even in principle, due to the complexity (closedness) of algorithms for processing large
 volumes of data. This is especially important when it comes to scientific knowledge, especially mathematical
 knowledge, which is needed in a wide range of classical and applied problems. In mathematical subject
 areas, it is important to rely on proven sources, to identify LLM hallucinations from true search results.

036 Some search engines that use LLM provide source references. The most professional and sophisticated 037 product is *Perplexity AI* (https://www.perplexity.ai). *Perplexity AI* allows a question, including in Russian, 038 and gives an answer in English and Russian with a link to Wikipedia and articles from mathnet, matem.anrb, *diffiournal.spbu* and other mathematical resources on the Internet. This version of the search engine, if we do not talk about links to Wikipedia, certainly serves as an example of checking the authenticity (factuality) 040 of the LLM answer. Naturally, it does not assess the correctness of the answer, from a mathematical point of 041 view, but it does provide a link to the publication, which, as a rule, is the purpose of the search. The question 042 of the evidence of mathematical conclusions Zong & Krishnamachari (2023) is not considered by us in this 043 work.

However, despite the perfection of *Perplexity AI*, which works with the entire Internet space, there remains the problem of working with special requests in a specific subject area, in particular in Russian, where clarification is required, and not general information, which, in principle, the user can find himself, spending, of course, more time than with *Perplexity AI*.

As a rule, for scientific fields (and mathematics, of course), it is necessary to analyze a specific collection of articles, including archived ones, full texts of which are not in the public domain. These articles need to be collected in a digital library, processed, and only then can the result come to the attention of LLM. To extract knowledge from these texts, previously not found in the public domain, it is necessary to provide them for LLM training in a new subject area, i.e. to compile a corpus of articles and describe this set semantically. This endless process is still relevant, since new subject areas appear, and new interdisciplinary studies with new terminology are added to traditional ones, which means that language models need to be adapted.

In this paper, the problem of LLM adaptation is proposed to be solved by using the knowledge graph (KG)
 MathSemanticLib of the semantic library *LibMeta* Olga Ataeva & Tuchkova (2024b), bypassing which,
 LLM extracts the answer from the subject area of mathematics and its application. The result is achieved due to the fact that the KG represents structured data, relies on the ontology and thesaurus of the subject area.

The structure of the article is as follows: Introduction, Related works, Data model of the Kg MathSemanticLib, Supervised knowledge extraction example LLM answer and Conclusion.

063 064 065

066

062

2 RELATED WORKS

067 The issues of joint consideration of the problems of constructing KG and LLM arose naturally, as a 068 continuation of the ideas of providing access to knowledge as structured data. The Awesome-LLM-KG 069 (https://github.com/RManLuo/Awesome-LLM-KG) page presents a collection of links to papers and resources about unifying LLMs and KGs. It graphically displays the advantages and disadvantages of LLMs 070 and KGs in the context of their mutual complementarity. The main idea is that the KG structure contributes 071 to improving LLM reasoning, and the linguistic capabilities and generalizations of LLMs improve the un-072 derstanding of the essence of knowledge in the KG. Awesome-LLM-KG also provides generalizations of 073 research directions and applications of the results of unifying LLMs and KGs in searching, building dia-074 log systems and AI assistants, and research methods. The synergetic nature of unifying LLMs and KGs is 075 separately noted, which is based on the mutual enrichment of LLMs and KGs when they are combined. 076

In the work Shirui Pan (2024) a cyclic procedure of integration of the domain knowledge and LLM is considered, as a result of which the LLM response and the domain knowledge itself are corrected, which is closest to the idea of our research.

The authors Linhao Luo (2024) provide an overview of the weaknesses of LLM related to hallucinations.
The authors see an improvement in the quality of inference in the use of KG for training LLM, but they note that this process is quite complex, since first a full KG must be constructed, and then LLM reasoning with graph constraints. The work Linhao Luo (2024) proposes a procedure for transforming KG for further traversal into LLM and for generating correct reasoning paths.

The option of creating a semantic description of mathematical concepts from school to university is considered in the work Samuel Debray (2025). Here Samuel Debray (2025) provides an overview of research related to the attempt to reflect the process of cognition of mathematical subject areas and their reflection in digital resources. This idea itself has haunted the scientific community, starting with the GDML project Patrick Ion & Zheng (2019). The research Samuel Debray (2025) uses the *GloVe* (https://nlp.stanford.edu/projects/glove/) algorithm on a large corpus of mathematical texts to identify the frequency of use of terms and their relationships.

The comparison is made between Wikipedia terms in French and their English translations, and the use of words from the dictionary by participants (https://osf.io/dxg2w) with different mathematical backgrounds. In this way, the *GloVe* algorithms were tested and a relatively good correspondence between the *GloVe* vectors and human judgments was established. This study Samuel Debray (2025) is important for our discussions in terms of the participation of experts in assessing the results of the semantic representation of subject areas and the reflection of these representations in the processes of cognition of mathematical areas. Like the authors of Samuel Debray (2025), we use the opinion of experts, but when creating semantic images of subject areas, we rely on classical sources such as encyclopedias and monographs.

100 The research Ruiqing Ding (2023) is devoted to the description of subject areas. The data are given on 101 how individual examples trained on corpora of specific subject areas achieve good results. However, it is 102 noted that this is not enough to create a general approach for different subject areas. The authors propose the 103 *KnowledgeDA* tool, a unified domain language model development service that can automatically generate a 104 domain language model by performing three steps: (i) localize domain knowledge entities in texts using an embedding-similarity approach; (ii) generate enriched samples by extracting exchangeable pairs of domain 105 entities from two representations of both the knowledge graph and the training data; (iii) select high-quality 106 enriched samples for fine-tuning using confidence-based scoring. 107

A *KnowledgeDA* prototype for learning language models for two domains: healthcare and software development. This example of creating text corpora by subject area is quite problematic to extend to mathematical subject areas, since the original sources may differ radically in structure and presentation features (for example, the presence of formulas changes the process of text preprocessing).

The work Iz Beltagy (2019) is devoted to training the *SciBERT* (https://github.com/allenai/scibert/) model on scientific texts, where the possibilities of improving BERT after unsupervised pre-training on a large multi-domain corpus of scientific publications are demonstrated. The *BERT* model architecture Jacob Devlin (2019) is based on a multilayer bidirectional Transformer Ashish Vaswani & Polosukhin (2017) is used.

Our work presents a technology for constructing KG, starting from arrays of texts of scientific mathematical and interdisciplinary journals, to the integration of the obtained KG *MathSemanticLib* with LLM in the journal recommendation system in the environment of the semantic library *LibMeta* Olga Ataeva & Tuchkova (2022).

121 122

123 124

3 DATA MODEL OF THE KG MathSemanticLib

The approach used in this work is that first an ontology and thesaurus of the subject area are built, then a KG based on the ontology, and then LLM is used for communication in the library. The data structure and ontology model of the *LibMeta* library for the KG *MathSemanticLib* are described in the works Olga Ataeva & Tuchkova (2023a), Olga Ataeva & Tuchkova (2024a), here we will note only some of their properties, namely: integration of various sources (encyclopedias, monographs, journals, classifiers, thesauri, dictionaries, formulas) based on the ontology; construction of a KG based on the ontology; use of a KG for organizing a dialogue in the library.

Thesauri contain the main terms of LibMeta subject areas, linked by hierarchical and horizontal relationships. The data model in LibMeta is an ontology in OWL (which is represented as an RDF graph). Filling the library is a process of completing the ontology by integrating data in accordance with their descriptions and metadata. The subject area is defined by forming a thematic subspace in the library ontology and establishing semantic links with the basic content of the library Olga Ataeva & Tuchkova (2023b).

The mathematical encyclopedia [ME] [EM], the encyclopedia of mathematical physics [MathPh], the the saurus of ordinary differential equations, the dictionary of special functions of mathematical physics and
 other Russian-language sources and components of the library [Ataeva3] are used as external basic tax onomies with which publications are linked. The creation and development of the LibMeta library [Ataeva0]

141 is based on the integration of mathematical knowledge, both in the retrospective and prospective direction, 142 by adding publications from various new subject areas of mathematics, related sciences and applications. 143 144 145 146 3.1 ONTOLOGY 147 148 149 150 The LibMeta digital library ontology defines the data structure. The concepts that make up the LibMeta 151 ontology are conventionally divided into concepts intended for: 152 153 - describing the content of a subject area; 154 - forming a thesaurus of any subject area; 155 156 - describing thematic collections; 157 – describing the task of integrating library content with source data from LOD. 158 159 Semantically significant connections are defined between these groups of concepts. The following formal definitions are used to describe the ontology: 160 161 Definition 1. Library thesaurus $TH = \{T, R\}$, where T are terms and R are the relationships between them. 162 Definition 2. Library content $C = \{IR, A, IO\}$, where IR are types of information resources, a set of 163 attributes $A \{a_i\}$, information objects $\{IO\}$. 164 Definition 3. Semantic labels $M = \{m_i\}$ of an information object are terms that are not included in the 165 thesaurus, but are necessary for thematic division of information objects IO within the subject area. 166 167 Definition 4. Semantically significant relationships of the library $P = \{P_i\}$ are the following main relation-168 ships: 169 $P_1(t, io)$ thesaurus term \rightarrow information object; 170 171 $P_2(io, t)$ information object \rightarrow thesaurus term; 172 $P_3(r, s)$ information resource \rightarrow class of source objects, where information resource is a general definition 173 for information objects stored in the system; thus, in fact, information objects are instances of information 174 resources: 175 $P_4(a, sa)$ information resource attribute \rightarrow property of source class; 176 177 $P_5(io, os)$ information object \rightarrow instance of class from data source; 178 $P_6(m, io)$ semantic label \rightarrow information object; 179 $P_7(io, m)$ information object \rightarrow semantic label. 180 181 In fact, the concepts are divided into three categories: the first includes definitions of the concepts of the 182 semantic library content, the second category refers to the definition of concepts necessary to support terms 183 in the thesaurus of the subject area, and the third includes definitions necessary to define the processes 184 of integrating the content of these resources. Based on these definitions, the main processes are described, 185 such as, for example, integrating data from different sources, categorization/classification, mapping different models of source data to a given subject area, constructing equivalence classes, etc. Fig. 1 shows the diagram of the article's links in the LibMeta ontology. 187





The construction of KG *MathSemanticLib* can be described in two global stages. At the first stage, a 'zero' version of KG is constructed from some source, and at the 'second' stage, the integration of the graph of in-coming data with the general graph of the library occurs by establishing links with the thesaurus Olga Ataeva & Tuchkova (2024b). The 'zero' version of the graph KG *MathSemanticLib* is the KG of the mathematical encyclopedia Vinogradov (1977–1985), EM (2022), and the 'second' stage is the integration of an array of scientific articles. When completing the ontology, KG is also completed, that is, the 'second' stage is each subsequent stage.

The main stages of data processing for GC are closely related to the sources from which the data comes.
Often the data is presented in an unstructured or semi-structured form. In our case, we consider, among
other things, unstructured texts of Russian-language scientific articles. Nodes can be larger ontology objects
Fig. 4, or objects - publication, term, person, formula.





- the formula is a semantic object in *LibMeta*, it can also be a KG node, for example, Fig. 5, which shows the links of the formula with other KG objects of the *MathSemanticLib* library.
- 379 There are several types of work with formulas in the *LibMeta* library:
- selection of basic formulas related to concepts;
- selection of formulas related to the concept;
 - selection of all formula links.
- 383 384 385 386

397

4 SUPERVISED KNOWLEDGE EXTRACTION EXAMPLE LLM ANSWER

Diving into the terminology of the subject area means accepting some non-standard terms established by
 specialists. For mathematical subject areas, it is important to reflect the meaning of formulas and terms
 when describing subject areas in digital libraries. Fig. 4–6 shows fragments of the *MathSemanticLib* KG for
 a concept, formula, classifiers.

Below is a comparison of responses from some language models.

ChatGpt 4 and YandexGPT 4 Pro (https://console.yandex.cloud/) coped with the query 'Lame equation',
 but gave only general information without links and formulas with an error, *Perplexity* and *LLAMA* (Large Language Model Meta AI https://app.hyperbolic.xyz/models/llama-3-3-70b-instruct) did not cope at all (Fig. 7-9).

398					
399	уравнение	Ламе		Search Image	
400			-	E Search Videc	
401	UL CARGE PARTA Antonio Antonio Antonio Antonio Destanto Antonio Martino Antonio	un en el construcción de la construcción de		Generate Ima	
402	No. a	image.jpg			
403	p Answer		TRUCT TOLL US HOUSE D ROUCKODOŬ		
404	соласно полсковол выдаче предлагают поведение ламы:	г выдаче, уравнения ламе не сущес ся следующие варианты уравнений	пвует. тем не менее, в поисковой й, которые могли бы описать		
405	1. Уравнение движ	кения ламы:			
406	$dx/dt = v * \cos(\theta)$	θ) dy/dt = v * sin(θ)			
407	где x и y - коорд	динаты ламы, v - скорость ламы, θ -	- угол движения ламы.		
408	2. Уравнение энер	гии ламы:			
409	$E = m * v^2/2 + r$	m * g * h			
410	где Е - энергия Л падения, h - выс	ламы, m - масса ламы, v - скорость сота ламы над уровнем земли.	ламы, g - ускорение свободного		
411	3. Уравнение пове	дения ламы:			
410	P = (F + H) / (S +	+ A)			
412	где Р - вероятно	ость того, что лама будет вести себ	я спокойно, F - количество		
413	корма, которое з S - количество с	стресса, которое лама испытывает,	мания, которое лама получает, А - количество атак, которое		
414	лама испытывае	РТ.			
415	A Share 🗋 Export	t ⊈ Rewrite	4 Q		
416	Ask follow-u	qu			
417					
418					
419	Figure 8: Pern	<i>lexity</i> answers to f	he query 'Lame equ	uation'.	
420					
421	The language model + knowledge grar	h allow using the	generalizing capah	ilities of the langua	age model to
422	explore the domain graph in natural la	inguage. Ouestion	ns formulated in na	tural language are	transformed

4	
6	Уравнение Ламе
0	
7	
3	оболочек. Оно используется в механике твёрдого тела для анализа распределения
9	напряжений и деформаций в этих объектах.
0	Уравнение Паме имеет спелующий вил:
-	Paulienne Marie Mineer energineum ung.
2	$\frac{d(r\sigma_{rr})}{dr} + \frac{\sigma_{rr}}{\sigma_{0}} = 0$
2	где \$σ_{rr}\$ — радиальное напряжение, а \$σ_{00}\$ — окружное (тангенциальное)
л	напряжение. Это уравнение позволяет определить распределение напряжений в
*	зависимости от расстояния от центра трубы или оболочки.
	Для решения уравнения Ламе необходимо задать граничные условия, которые
2	определяют напряжения на внутренней и внешней поверхностях трубы или оболочки. После этого можно использовать различные метолы решения лифференциальных
r 2	уравнений, такие как метод разделения переменных, метод конечных разностей или
5	метод конечных элементов.
9	- С С Скопировать ответ
0	
1	
2	
3	
4	
5	Figure 9: <i>Yandex</i> answers to the query 'Lame equation'.
6	
7 into SP	AROL queries based on the ontological model of the graph. The model's answers are given based
⁸ the mat	hematical encyclonedia articles and their connections Fig. 10
9	IN ADALMAD AN AVAILABATICA ALLEAN ADDITING A ANDRA ADDIN 1219 - 117
The mo	nematical encyclopedia articles and then connections Fig. 10.
the-arf	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-or
	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi
2 models	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi. They are text-to-text, decoder-only large language models, available in English, with open weigh
models for both	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state- open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety
models for both text gen	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety iteration tasks, including question answering, summarization, and reasoning. Their relatively smallers
models for both text get size ma	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weight pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smalkes it possible to deploy them in environments with limited resources such as a laptop, deskty own cloud infrastructure, democratizing access to state of the art AI models and helping fost
models for both text ger size ma or your	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, desktor own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://buggingface.co/google/gemma-2-9b-it.)
models for both text get size ma or your innovat	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state- open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh n pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smalkes it possible to deploy them in environments with limited resources such as a laptop, deskto own cloud infrastructure, democratizing access to state of the art AI models and helping fost ion for everyone (https://huggingface.co/google/gemma-2-9b-it).
models for both text get size ma or your movat	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it).
models for both text gen size ma or your innovat	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state- open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh or pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety peration tasks, including question answering, summarization, and reasoning. Their relatively smitkes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fost ion for everyone (https://huggingface.co/google/gemma-2-9b-it).
models for both text gen size ma or your innovat	 del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemi . They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fost ion for everyone (https://huggingface.co/google/gemma-2-9b-it). m 1: How many publications?
models for both text get size ma or your innovat	del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemma. They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it).
models for both text gen size ma or your nnovat Questic LECT	 del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gemia. They are text-to-text, decoder-only large language models, available in English, with open weight pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, desktor own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it). on 1: How many publications? Generated SPARQL: SE COUNT(*) AS ?publications)
models for both text get size ma or your innovat Questic LECT Ouestic	 and there connections Fig. 10. and there connections Fig. 10. and encyclopedia and en
models for both text get size ma or your innovat Questic LECT Questic	 adel <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gem. They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it). b Generated SPARQL: SE COUNT(*) AS ?publications) m 2: How many concepts?
models for both text get size ma or your mnovat Questic LECT Questic	 adel <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gem. They are text-to-text, decoder-only large language models, available in English, with open weigh pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety beration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fossion for everyone (https://huggingface.co/google/gemma-2-9b-it). on 1: How many publications? Generated SPARQL: SE COUNT(*) AS ?publications) on 2: How many concepts? SELECT (COUNT(DISTINCT ?concept) AS ?count)
models for both text get size ma or your nnovat Questic LECT Questic	 del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gem. They are text-to-text, decoder-only large language models, available in English, with open weigh pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, deskter own cloud infrastructure, democratizing access to state of the art AI models and helping fossion for everyone (https://huggingface.co/google/gemma-2-9b-it). on 1: How many publications? Generated SPARQL: SE COUNT(*) AS ?publications) m 2: How many concepts? SELECT (COUNT(DISTINCT ?concept) AS ?count) WHERE
models for both text gen size ma or your innovat Questic Questic	 del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gem. They are text-to-text, decoder-only large language models, available in English, with open weigh pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety teration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, desktor own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it). m 1: How many publications? Generated SPARQL: SE COUNT(*) AS ?publications) m 2: How many concepts? SELECT (COUNT(DISTINCT ?concept) AS ?count) WHERE ?concept a :Concept .
models for both text gen size ma or your innovat Questic LECT of Questic	 del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gem. They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smakes it possible to deploy them in environments with limited resources such as a laptop, desktor own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it). on 1: How many publications? Generated SPARQL: SE COUNT(*) AS ?publications) on 2: How many concepts? SELECT (COUNT(DISTINCT ?concept) AS ?count) WHERE ?concept a :Concept .
models for both text get size ma or your innovat Questic Questic Questic	 del <i>Gemma2-9b-It</i> was used with following 3 questions. Gemma is a family of lightweight, state-open models from Google, built from the same research and technology used to create the Gem. They are text-to-text, decoder-only large language models, available in English, with open weigh a pre-trained variants and instruction-tuned variants. Gemma models are well-suited for a variety heration tasks, including question answering, summarization, and reasoning. Their relatively smukes it possible to deploy them in environments with limited resources such as a laptop, deskto own cloud infrastructure, democratizing access to state of the art AI models and helping fostion for everyone (https://huggingface.co/google/gemma-2-9b-it). m 1: How many publications? Generated SPARQL: SE COUNT(*) AS ?publications) m 2: How many concepts? SELECT (COUNT(DISTINCT ?concept) AS ?count) WHERE ?concept a :Concept . m 3: How many concepts are connected with Lame formula?

470	
471	
472	Llama-3.3-70b-Versatile 🗘 🛛 R
473	
474	Try it
475	Ламе уравнение
470	
478	зравнения ламе: это математическое уравнение, которое описывает поведения ламы, но на самом деле это не существует. Однако, я могу предложить вам несколько вариантов уравнений, которые могли бы описать поведение
479	
480	$dx/dt = v * \cos(\theta) dy/dt = v * \sin(\theta)$
481	где х и у - координаты ламы, v - скорость ламы, θ - угол движения ламы.
482	2. Уравнение энергии ламы:
483	E = m * v^2 / 2 + m * g * h
484	где Е - энергия ламы, m - масса ламы, v - скорость ламы, g - ускорение свободного падения, h - высота ламы над уровнем земли.
485	3. Уравнение поведения ламы:
486	P=(F+H)/(S+A)
487	где Р - вероятность того, что лама будет вести себя спокойно, F - количес€Фво корма, которое лама получает, Н - иопищество рымарция имполов прака полицает S - иопищество сторого волка имполов прака имполовета.
488	-
489	
490	
491	
492	Figure 10: LLAMA Answers to the Query 'Lame Equation'.
493	
494	• SELECT (COUNT(DISTINCT ?concept) AS ?count)
495	WHERE
497	Pformula a : LameFormula
498	Proncent a Concent
499	2 Some pt a Concept .
500	Stores of conceptions a SDA DOL success and monopole to it using a CDT like model
501	Stages of generating a SPARQL query and response to it using a GPT like model.
502	Stage 1. Translation of text into SPARQL query.
503	General format of the prompt for a model to generate a SPARQL query includes:
504	Question (promt): in natural language
505	Scheme: The graph data model is specified as an rdfs scheme, which is annotated with the rdfs:label
506	and rdfs:comment properties, which contain a natural language description of the classes and prop-
507	erties of the ontology
508	Instructions:
509	- Generate a SPARQL SELECT query to query a graph database using the following ontology
510	schema in Turtle format: {schema}
510	 Use only the classes and properties specified in the schema.
512	 Do not include any that are not explicitly provided.
512	- Ensure that all required prefixes are included.
515	- Output only the SPARQL query without any backticks or additional text.
516	- The question is: {prompt}
- · ·	The American In: (how he)

517	Frample 1
518	Ouaction (nomt): in natural laguage How many concents are there?
519	Salarma (from t). In natural laguage How many concepts are there?
520	Scheme (Iragment):
521	
522	http://libmeta.ru/thesaurus/concept/DE0002
523	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
524	http://libmeta.ru/Concept .
525	http://libmeta.ru/thesaurus/ODU
526	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
527	http://libmeta.ru/Thesaurus.
528	· · · · · · · · · · · · · · · · · · ·
529	Answer
530	PREFIX rdfs: http://www.w3.org/2000/01/rdf-schema#
531	PREFIX rdf http://www.w3.org/1999/02/22-rdf-syntax_ns#
532	SELECT (COUNT(2concent) AS 2concentCount)
533	WHEDE
534	
535	/concept rdf:type inttp://ilometa.ru/Concept¿.
536	Stage 2. Graph query
537	Next comes the step of sending the generated query through the SPARQL access point and receiving
538	the response. For the query specified in the example, the response comes in the form
539 540	conceptCount': 'datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal', 'value': '1221'
541	Stage 3. Translate the answer into humanreadable format
542	General format of the prompt to translate the response into humanreadable format includes:
543	Ouestion (promt): in natural language
544	Question (promit). In natural language
545	and rdfs: comment properties, which contain a natural language description of the classes and prop-
546	erties of the ontology
547	Instructions:
548	Create a natural language meaning have dealake on the merits of a SDADOL success
549	- Create a natural language response based solely on the results of a SPARQL query.
550	- You are an assistant who produces clear, human-friendly answers.
551	- Use only the provided information to build your response.
552	 This information is authoritative
553	 Do not question or alter it with your own knowledge.
554	 Ensure your answer reflects an AI assistants tone without adding any extra details.
555	 If no information is available, simply state that you don't know.
556	– Schema: {schema}
557	- Context: {context}
558	- Ouestion: {prompt}
559	Example 2
560	Crucition (promt): in natural laguage How many concents are there?
561	Question (promit): in natural laguage from many concepts are there:
562	
563	http://libmeta.ru/thesaurus/concept/DE0002

 http://libmeta.ru/Concept. http://libmeta.ru/Concept. http://libmeta.ru/Concept. http://libmeta.ru/Thesaurus. Context: {'conceptCount': {datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal', 'value': '1221'} Answer Answer There are 1.221 concepts in total. Stage 4. Correction of the query If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: Question (promt): in natural laguage Scheme (fragment): http://libmeta.ru/thesaurus/CODU http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/DDU http://libmeta.ru/thesaurus/EVE The error is: {error.message} The error is: {error.message} The error is: {error.message} Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. The error is: in natural laguage How many concepts are there? Scheme (fragment): The tron is: natural laguage How many concepts are there? Scheme (fragment): http://libmeta.ru/thesaurus/CDU http://libmeta.ru/thesaurus/CDU http://libmeta.ru/thesaurus/CDU 	564	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
 http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/ODU Stage 4. Correction of the query There are 1,221 concepts in total. Stage 4. Correction of the query If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: Question (promit): in natural laguage Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus/CODU http://libmeta.ru/thesaurus. Error: Error text SPARQI. wrong query Instruction - The error is: {error.message} - Provide a corrected version of the SPARQL query without altering its logic. - Include only the corrected query. No oreplanations or additional text. - Do not enclose the query in backticks. - The entology schema in Turtle format is provided below: {schema} Example 3. Question (fromt): in natural laguage How many concepts are there? Scheme (fragment): 	565	http://libmeta.ru/Concent
667 Intp://www.w3.org/1990/02/22-rdf-syntax-ns#type 668 http://ibmeta.ru/Thesaurus. Context: 710 {'conceptCount': { datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal', 'value': 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'walue': 'literal', 'literal', 'literal', 'literal', 'walue': 'literal', 'literal', 'walue': 'literal', 'literal', 'literal', 'literal', 'literal', 'literal', 'literal', 'literal', 'literal', 'walue': 'literal', '	566	http://libmata.ru/Concept .
 http://libmeta.ru/Thesaurus. Context: {'concept/Count': {'datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal', 'value': '1221'}} Answer There are 1,221 concepts in total. Stage 4. Correction of the query If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: Question (promt): in natural laguage Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/Concept/DE0012 http://libmeta.ru/thesaurus/Concept/DE0012 http://libmeta.ru/thesaurus/Concept. http://libmeta.ru/thesaurus/Concept. http://libmeta.ru/thesaurus/CODU http://libmeta.ru/thesaurus/ODU http://libmeta.ru/thesaurus . Error: Error text SPARQL wrong query Instruction The error is: {error.message} Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. Ne explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): tut://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/Concept. 	567	http://itometa.tu/itiesautus/ODO
 http://libmeta.ru/libesaurus. Context: {'conceptCount: {'datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal', 'value': '1221'}} Answer There are 1,221 concepts in total. Stage 4. Correction of the query If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: Question (promt): in natural laguage Scheme (fragment): 	568	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
 for ('conceptCount': {'datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal', 'yalue': 'literal', 'yalue': 'literal', 'alue': 'literal', 'yalue': 'literal', 'alue': 'literal', 'yalue': 'literal', 'literal', 'literal', 'literal', 'literal', 'literal', 'yalue': 'literal', 'liter	569	http://libmeta.ru/Thesaurus. Context:
571 value : 1221 }} Answer There are 1,221 concepts in total. 573 There are 1,221 concepts in total. 574 Stage 4. Correction of the query 575 If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: 577 Question (promt): in natural laguage 58 Scheme (fragment):	570	{'conceptCount': {'datatype': 'http://www.w3.org/2001/XMLSchema#integer', 'type': 'literal',
572 Answer 573 There are 1,221 concepts in total. 574 Stage 4. Correction of the query 575 If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: 576 Question (promt): in natural laguage 578 Scheme (fragment): 579 571 matural laguage 572 Scheme (fragment): 573 574 http://libmeta.ru/thesaurus/concept/DE0002 575 http://libmeta.ru/Concept . 576 http://libmeta.ru/thesaurus.ODU 577 Error text 578 Error text 579 Wrong query 571 Error text 572 Provide a corrected version of the SPARQL query without altering its logic. 573 - The serror is: {error_message} 574 - Provide a corrected query. 575 - The ontology schema in Turtle format is provided below: {schema} 576 - The ontology schema in Turtle format is provided below: {schema} 577 - The ontology schema in Turtle format is provided below: {schema} <td< td=""><td>571</td><td>'value': '1221'}}</td></td<>	571	'value': '1221'}}
573 There are 1,221 concepts in total. 574 Stage 4. Correction of the query 575 If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: 577 Question (promt): in natural laguage 578 Scheme (fragment): 579 580 580 581 http://libmeta.ru/thesaurus/concept/DE0002 582 http://libmeta.ru/thesaurus/concept/DE0012 583 http://libmeta.ru/Concept . 584 http://libmeta.ru/Concept . 584 http://libmeta.ru/Thesaurus / Error: 587 Error text 588 SPARQL 589 wrong query 580 Instruction 591 - The SPARQL query below 592 - The error is: {error.message} 593 - The error is: {error.message} 594 - No explanations or additional text. 595 - No explanations or additional text. 596 - The ontology schema in Turtle format is provided below: {schema} 597 - The ontology schema in Turtle format is provided	572	Answer
574 Stage 4. Correction of the query 575 If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: 577 Question (promt): in natural laguage 578 Scheme (fragment): 579 580 581 http://libmeta.ru/thesaurus/concept/DE0002 582 http://libmeta.ru/thesaurus/concept/DE0002 583 http://libmeta.ru/thesaurus/DDU 584 http://libmeta.ru/thesaurus/DDU 585 http://libmeta.ru/Thesaurus/DDU 586 http://libmeta.ru/Thesaurus. Error: 587 Error text 588 SPARQL 589 wrong query 580 Instruction 581 is invalid. 582 - The SPARQL query below (generated.sparql) is invalid. 584 - The error is: {error_message} - Frevide a corrected query. - No explanations or additional text. - Do not enclose the query in backticks. - The ontology schema in Turtle format is provided below: {schema} Example 3. Question (promt): in natural laguage How many concepts are there? Scheme (fragment): - h	573	There are 1,221 concepts in total.
575 If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions: 577 Question (pront): in natural laguage 578 Scheme (fragment): 579 571 http://libmeta.ru/thesaurus/concept/DE0002 572 http://www.w3.org/1999/02/22-rdf-syntax-ns#type 573 http://libmeta.ru/thesaurus/ODU 574 http://libmeta.ru/thesaurus/ODU 575 Http://libmeta.ru/Thesaurus . Error: 576 Http://libmeta.ru/Thesaurus . Error: 577 Error text 588 SPARQL 590 wrong query 591 Instruction 592 - The SPARQL query below {generated.sparql} 593 is invalid. 594 - The error is: {error.message} 595 - Include only the corrected query. 596 - Provide a corrected version of the SPARQL query without altering its logic. 597 - No explanations or additional text. 598 - Do not enclose the query in backticks. 599 - No explanations or additional text. 590 - Do not enclose the query in backticks.	574	Stage 4. Correction of the query
977 Question (pront): in natural laguage 978 Scheme (fragment): 979 971 http://libmeta.ru/thesaurus/concept/DE0002 972 http://libmeta.ru/Concept . 973 http://libmeta.ru/Concept . 974 http://libmeta.ru/Concept . 975 http://libmeta.ru/Concept . 976 http://libmeta.ru/Concept . 977 Entror two.sorg/1999/02/22-rdf-syntax-ns#type 978 http://libmeta.ru/Thesaurus/ODU 979 Instruction 970 Error text 971 Error text 972 - The SPARQL query below {generated.sparql} 973 is invalid. 974 is invalid. 975 - The error is: {error.message} 976 - The error is: {error.message} 977 - No explanations or additional text. 978 - Do not enclose the query in backticks. 979 - The ontology schema in Turtle format is provided below: {schema} 978 - The ontology schema in Turtle format is provided below: {schema} 979 - The ontology schema in Turtle format is provided below: {schema}	575 576	If an invalid request was generated in response to the first step, an attempt is made to correct it using the following instructions:
Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/Concept. http://libmeta.ru/Concept. http://libmeta.ru/Concept. http://libmeta.ru/Concept. http://libmeta.ru/thesaurus/ODU http://libmeta.ru/Toksaurus. Error Barton (ragment): Barton (ragment):	577	Question (promt): in natural laguage
579 561 http://libmeta.ru/thesaurus/concept/DE0002 561 http://libmeta.ru/Concept . 562 http://libmeta.ru/Concept . 563 http://libmeta.ru/Concept . 564 http://libmeta.ru/thesaurus/ODU 565 http://libmeta.ru/Thesaurus . 566 http://libmeta.ru/Thesaurus . 567 Error text 568 SPARQL 579 Wrong query 581 Instruction 592 - The SPARQL query below 593 {generated sparql} 594 is invalid. 595 - The error is: {error_message} 596 - The error is: {error_message} 597 - Provide a corrected query. 598 - No explanations or additional text. 599 - No explanations or additional text. 590 - The ontology schema in Turtle format is provided below: {schema} 591 Example 3. 592 - The ontology schema in Turtle format is provided below: {schema} 593 - The ontology schema in Turtle format is provided below: {schema} 594 Example 3.	578	Scheme (fragment):
impose impose impose impose <td< td=""><td>579</td><td>Scheme (mugment)</td></td<>	579	Scheme (mugment)
antp://initiation/initiation/oncept/DE0002 http://iwww.w3.org/1999/02/22-rdf-syntax-ns#type http://libmeta.ru/Concept . http://libmeta.ru/Thesaurus/ODU http://libmeta.ru/Thesaurus . Error text SPARQL wrong query Instruction -	580	 http://libmata.m/theseume/concent/DE0002
http://www.w3.org/1999/02/22-rdf-syntax-ns#type http://libmeta.ru/Concept http://libmeta.ru/Thesaurus/ODU http://libmeta.ru/Thesaurus . Error: Error text SPARQL wrong query Instruction -	581	http://itometa.ru/thesaurus/concept/DE0002
583 http://libmeta.ru/thesaurus/ODU 584 http://libmeta.ru/thesaurus/. Error: 585 http://libmeta.ru/Thesaurus . Error: 586 SPARQL 588 SPARQL 589 wrong query 590 Instruction 591 - The SPARQL query below {generated_sparql} is invalid. 592 - The error is: {error_message} 593 - Include only the corrected query without altering its logic. 595 - Include only the corrected query. 596 - No explanations or additional text. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/thesaurus/concept/DE002 609 http://libmeta.ru/thesaurus/concept. 609 http://libmeta.ru/thesaurus/CODU	582	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
584 http://libmeta.ru/thesaurus/ODU 585 http://libmeta.ru/Thesaurus / Error: 586 http://libmeta.ru/Thesaurus / Error: 587 Error text 588 SPARQL 589 wrong query 590 Instruction 591 - The SPARQL query below {generated_sparql} is invalid. 592 - The error is: {error_message} 593 - The error is: {error_message} 594 - The error is: {error_message} 595 - Include only the corrected query. 596 - No explanations or additional text. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/thesaurus/concept/DE0002 609 http://libmeta.ru/thesaurus/concept. 609 http://libmeta.ru/thesaurus/concept.	583	http://libmeta.ru/Concept.
585 http://www.w3.org/1999/02/22-rdf-syntax-ns#type 586 http://libmeta.ru/Thesaurus . Error: 587 Error text 588 SPARQL 589 wrong query 590 Instruction 591 - The SPARQL query below {generated.sparql} is invalid. 593 - The error is: {error_message} 594 - The error is: {error_message} 595 - The error is: {error_message} 596 - No explanations or additional text. 599 - No explanations or additional text. 590 - The ontology schema in Turtle format is provided below: {schema} 592 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/Concept . 609 http://libmeta.ru/Concept . 609 http://libmeta.ru/Concept . 609 http://libmeta.ru/Concept .	584	http://libmeta.ru/thesaurus/ODU
586 http://libmeta.ru/Thesaurus . Error: 587 Error text 588 SPARQL 589 wrong query 590 Instruction 591 - The SPARQL query below {generated_sparql} is invalid. 593 - The error is: {error_message} 594 - The error is: {error_message} 595 - Provide a corrected version of the SPARQL query without altering its logic. 596 - Include only the corrected query. 599 - No explanations or additional text. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/thesaurus/concept. 609 http://libmeta.ru/thesaurus/concept . 609 http://libmeta.ru/thesaurus/concept .	585	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
 587 Error text 588 SPARQL 589 wrong query 590 Instruction 591 - The SPARQL query below {generated_sparql} is invalid. 593 - The error is: {error_message} 596 - The error is: {error_message} 596 - Provide a corrected version of the SPARQL query without altering its logic. 596 - Include only the corrected query. 599 - No explanations or additional text. 590 - Do not enclose the query in backticks. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/Concept . 609 http://libmeta.ru/Concept . 	586	http://libmeta.ru/Thesaurus . Error:
 SPARQL wrong query Instruction The SPARQL query below {generated_sparql} is invalid. The error is: {error_message} The error is: {error_message} Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. No explanations or additional text. No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/concept . http://libmeta.ru/concept . 	587	Error text
 wrong query Instruction The SPARQL query below {generated_sparql} is invalid. is invalid. The error is: {error_message} The error is: {error_message} Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept. http://libmeta.ru/Concept . http://libmeta.ru/Concept . 	588	SPARQL
590 Instruction 591 - The SPARQL query below 592 - The SPARQL query below 593 - generated_sparql} 594 is invalid. 595 - The error is: {error_message} 596 - Provide a corrected version of the SPARQL query without altering its logic. 597 - Include only the corrected query. 598 - Include only the corrected query. 599 - No explanations or additional text. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/Concept . 609 http://libmeta.ru/thesaurus/ODU	589	wrong query
 591 The SPARQL query below {generated_sparql} is invalid. 595 The error is: {error_message} Provide a corrected version of the SPARQL query without altering its logic. 596 Provide a corrected query. Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/Concept . 609 http://libmeta.ru/thesaurus/ODU 	590	Instruction
 Interstructure query below {generated_sparql} is invalid. The error is: {error_message} Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/CODU http://libmeta.ru/thesaurus/ODU 	591	- The SPAROL query below
 is invalid. The error is: {error_message} Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept . http://libmeta.ru/thesaurus/ODU 	592	{generated sparal}
 594 595 - The error is: {error_message} 596 - Provide a corrected version of the SPARQL query without altering its logic. 598 - Include only the corrected query. 599 - No explanations or additional text. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://www.w3.org/1999/02/22-rdf-syntax-ns#type 609 http://libmeta.ru/thesaurus/ODU 	593	is invalid.
 595 - The error is: {error_message} 596 - Provide a corrected version of the SPARQL query without altering its logic. 598 - Include only the corrected query. 599 - No explanations or additional text. 600 - Do not enclose the query in backticks. 601 - The ontology schema in Turtle format is provided below: {schema} 602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://libmeta.ru/thesaurus/ODU 609 http://libmeta.ru/thesaurus/ODU 	594	
 Provide a corrected version of the SPARQL query without altering its logic. Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/concept . http://libmeta.ru/thesaurus/ODU 	595	- The error is: {error_message}
 Include only the corrected query. No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://www.w3.org/1999/02/22-rdf-syntax-ns#type http://libmeta.ru/thesaurus/ODU 	596	- Provide a corrected version of the SPAROL query without altering its logic.
 No explanations or additional text. Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept . http://libmeta.ru/thesaurus/ODU 	597	- Include only the corrected auery.
 by the explanations of definitional total Do not enclose the query in backticks. The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://libmeta.ru/thesaurus/concept . http://libmeta.ru/thesaurus/ODU 	598	 No explanations or additional text
 b) b) hot checkse the query in backfields. - The ontology schema in Turtle format is provided below: {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://www.w3.org/1999/02/22-rdf-syntax-ns#type http://libmeta.ru/thesaurus/ODU 	599	 Do not enclose the query in backticks
 File ontology schema in Turtle format is provided below. {schema} <i>Example 3.</i> Question (promt): in natural laguage How many concepts are there? Scheme (fragment): Scheme (fragment): http://libmeta.ru/thesaurus/concept/DE0002 http://www.w3.org/1999/02/22-rdf-syntax-ns#type http://libmeta.ru/Concept . http://libmeta.ru/thesaurus/ODU 	601	The optology scheme in Turtle format is provided below: [scheme]
602 Example 3. 603 Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://www.w3.org/1999/02/22-rdf-syntax-ns#type 609 http://libmeta.ru/Concept . 610 http://libmeta.ru/thesaurus/ODU	602	- The ontology schema in Turne format is provided below. {schema}
Output Question (promt): in natural laguage How many concepts are there? 604 Scheme (fragment): 605 606 607 http://libmeta.ru/thesaurus/concept/DE0002 608 http://www.w3.org/1999/02/22-rdf-syntax-ns#type 609 http://libmeta.ru/Concept . 610 http://libmeta.ru/thesaurus/ODU	603	Example 3.
Scheme (fragment): 605 606 607 608 609 609 609 600 601 602	604	Question (promt): in natural laguage How many concepts are there?
606607http://libmeta.ru/thesaurus/concept/DE0002608http://www.w3.org/1999/02/22-rdf-syntax-ns#type609http://libmeta.ru/Concept .610http://libmeta.ru/thesaurus/ODU	605	Scheme (fragment):
607http://libmeta.ru/thesaurus/concept/DE0002608http://www.w3.org/1999/02/22-rdf-syntax-ns#type609http://libmeta.ru/Concept .610http://libmeta.ru/thesaurus/ODU	606	
 http://www.w3.org/1999/02/22-rdf-syntax-ns#type http://libmeta.ru/Concept . http://libmeta.ru/thesaurus/ODU 	607	http://libmeta.ru/thesaurus/concept/DE0002
609 http://libmeta.ru/Concept . 610 http://libmeta.ru/thesaurus/ODU	608	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
610 http://libmeta.ru/thesaurus/ODU	609	http://libmeta.ru/Concept.
	610	http://libmeta.ru/thesaurus/ODU

611	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
612	http://libmeta.ru/Thesaurus
613	Fron.
614	400. Error MALEORMED OUEDV: ONome 'lhm: Theseurus Concent' uses on undefined marky
615	400: Error MALFORMED QUERY: QName Tom: ThesaurusConcept uses an underined prenx
616	SPARQL
617	PREFIX rdfs: http://www.w3.org/2000/01/rdf-schema#
618	PREFIX rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
619	SELECT (COUNT(?concept) AS ?conceptCount)
620	WHERE
621	?concept rdf:type lbm:ThesaurusConcept
622	Answer
623	DDEELY rdfs: http://www.w2.org/2000/01/rdf.schemott
624	PEFEV = 15 144 mm + 2 mm (1000/02/22 mm mm + 1000/02/22 mm + 10000/02/22 mm + 1000/02/22 mm +
625	PREFIX rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
626	PREFIX lbm: http://libmeta.ru/
627	SELECT (COUNT(?concept) AS ?conceptCount)
628	WHERE
629	?concept rdf:type lbm:Concept.
630	
631	
632	
633	Формула уравнение Ламе 🖉
634	Companying systemetry and Types
635	overlag at an annan ganta an annan ganta an
636	Copyra particle The second seco
637	Bortine Caudo 012 Program y presence Take Antipatrime con the State of
638	y person of the second se
639	Ventionen regenerationen regenerationen ventionen Hander ventionen Hander venti
640	
641	Kr. Terrera (Manader
642	
643	
644	
645	Figure 11: Graph for 'Lama equation' Math Somantial in VG and answer
646	Figure 11. Graph for Lame equation, <i>mainsemanticlub</i> KG and answer.
647	

Fig. 10 shows schematically T-T that the 'Lame equation' is selected and that it is linked to various objects of the library. These links arise because the formula in the *LibMeta* library is semantically linked through its context object.

5 CONCLUSION AND FINDINGS

The proposed approach to adapting LLM to a specific mathematical area of scientific Russian-language
 journals has been tested in the semantic digital library *LibMeta*. The experience of integrating LLM and the
 MathSemanticLib KG allows us to conclude that it is possible to organize a user dialogue with the library
 and create a digital assistant with the functions of a reader, author, editor, and reviewer of the journal.

It should be noted that when completing the *LibMeta* digital library ontology, it will be possible to ensure further integration of LLM and the *MathSemanticLib* KG, which means that the capabilities of a controlled Russian-language dialogue on mathematical topics will develop. At the same time, the ability to check the LLM output within the *LibMeta* content and external sources, which will continue to be integrated, will be preserved. Further research will continue in the direction of developing dialogue and recommender systems.

664 665 REFERENCES

663

672

680

681

682

683

698

- 666 Encyclopedia of Mathematics from Springer Verlag. Kluwer Academic Publishers, 2022.
- Niki Parmar Jakob Uszkoreit Llion Jones Aidan N. Gomez Lukasz Kaiser Ashish Vaswani, Noam Shazeer
 and Illia Polosukhin. Attention is all you need. *NIPS*, 2017.
- Alexandr Muromskii Eugenii Moiseev and Natalia Tuchkova. *Information search thesaurus of subject area Ordinary Differential Equations*', volume 1. MAKS Press, 2005.
- Ludwig Faddeev. *Matematicheskaya fizika*. *Enciklopediya*, volume 1. Bol'shaya Rossijskaya enciklopediya,
 1998.
- Arman Cohan Iz Beltagy, Kyle Lo. Scibert: A pretrained language model for scientific text. *Conference on Empirical Methods in Natural Language Processing.*, 2019.
- Kenton Lee Kristina Toutanova Jacob Devlin, Ming-Wei Chang. Bert: Pre-training of deep bidirectional
 transformers for language understanding. *NAACL-HLT*, 2019.
 - Reza Haf Chen Gong Shirui Pan Linhao Luo, Zicheng Zhao. Graph-constrained reasoning: Faithful reasoning on knowledge graphs with large language models. arXiv preprint arXiv:2410.13080, pp. 1–20, 2024.
- Vladimir Serebryakov Olga Ataeva and Natalia Tuchkova. Cifrovaya biblioteka po obyknovennym differ encial'nym uravneniyam na osnove libmeta. Nauchnyj servis v seti Internet: trudy XIX Vserossijskoj
 nauchnoj konferencii (18-23 sentyabrya 2017 g., g. Novorossijsk)., 44:21–33, 2017.
- Vladimir Serebryakov Olga Ataeva and Natalia Tuchkova. Development of the semantic space 'mathematics' by integrating a subspace of its applied area . *Lobachevskii Journal of Mathematics*, 43:29–40, 2022.
- Vladimir Serebryakov Olga Ataeva and Natalia Tuchkova. Ontological approach to a knowledge graph construction in a semantic library. *Lobachevskii Journal of Mathematics*, 44:2229–2239, 2023a.
- Vladimir Serebryakov Olga Ataeva and Natalia Tuchkova. From texts to knowledge graph in the semantic library libmeta. *Lobachevskii Journal of Mathematics*, 45:2211–2219, 2024a.
- Vladimir Serebryakov Olga Ataeva and Natalia Tuchkova. Ontology-driven knowledge graph construction
 in the mathematics semantic library. *Pattern Recognition and Image Analysis*, 34:451–458, 2024b.
- Vladimir Serebryakov Olga Ataeva, Yulia Kornet and Natalia Tuchkova. Approach to creating a thesaurus and a knowledge graph of an applied subject area. *Lobachevskii Journal of Mathematics*, 44:2577–2586, 2023b.
- Gadadhr Misra Alf Onshuus Stephen Watt Patrick Ion, Thierry Bouche and Liu Zheng. International mathematical knowledge trust imkt: An update on the global digital mathematics library. *Proceedings of the International Congress of Mathematicians (ICM 2018).*, pp. 1157–1175, 2019.

705 706 707	Anton Firc Ramil Malinka, Martin Pereni and all. On the educational impact of chatgpt: Is artificial intel- ligence ready to obtain a university degree? In Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2023), pp. 1–7, 2023.
708 709 710	Leye Wang Ruiqing Ding, Xiao Han. A unified knowledge graph augmentation service for boosting domain- specific nlp tasks. <i>Findings of the Association for Computational Linguistics</i> , pp. 353–369, 2023.
711 712 713	Stanislas Dehaene Samuel Debray. Mapping and modeling the semantic space of math concepts. <i>Cognition</i> , 254:1–8, 2025.
714 715 716	Yufei Wang Chen Chen Jiapu Wang Xindong Wu Shirui Pan, Linhao Luo. Unifying large language models and knowledge graphs: A roadmap. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 36:3580–3599, 2024.
717 718 719	Ivan Vinogradov. <i>Matematicheskaya enciklopediya (v 5 tomah)</i> , volume 1-5. Sovetskaya enciklopediya, 1977–1985.
720 721 722	Mingyu Zong and Bhaskar Krishnamachari. Solving math word problems concerning systems of equa- tions with gpt-3. <i>Proceedings of the Thirteenth AAAI Symposium on Educational Advances in Artificial</i> <i>Intelligence</i> , 37:15972–15979, 2023.
723	
724	
725	
727	
728	
729	
730	
731	
732	
733	
734	
735	
736	
737	
738	
739	
740	
741	
742	
743	
744	
745	
746	
747	
748	
749	
750	
101	