

ЦИПР-2019: Технологии хранения и анализа больших данных для цифровой экономики: новые результаты

Семантические сервисы обработки больших массивов научных документов и формирование цифровых библиотек

Елизаров А.М., Липачёв Е.К.

Центр хранения и анализа больших данных Научно-образовательного центра компетенций в области цифровой экономики МГУ

Казанский (Приволжский) федеральный университет

Иннополис. 22-24 мая 2019



Обсуждаемые вопросы

- Большие данные и научные коллекции
- Большие данные в математике
- Информационные технологии в математике
- Big Math & One-Brain Barrier
- Формализация математического знания
- Цифровые математические библиотеки
- Фабрика метаданных цифровой библиотеки
- Автоматизация издательских процессов

Большие данные в математике

Проблема четырёх красок

Appel K. and Haken W., *Every map is four Colourable* // Bulletin of the American Mathematical Society 82 (1976), 711–712.

Appel K., Haken W., *Every Planar Map is Four Colorable. I. Discharging* // Illinois J. of Mathematics, 21 (3), pp. 429–490 (1977).

Appel K., Haken W., Koch J., *Every Planar Map is Four Colorable. II. Reducibility* // Illinois Journal of Mathematics, 21 (3), pp. 491–567 (1977).

Проблема сведена к примерно 2000 конфигурациям, которые проверены компьютерной программой.

Правомерность компьютерных доказательств решается формализацией.

В 2005 году Д. Гонтиром теорема доказана формализацией на Coq.

Gonthier, Georges, *Formal Proof - The Four-Color Theorem* // Notices of the American Mathematical Society, 55 (11): 1382–1393 (2008)

Formal Proof—The Four-Color Theorem

Georges Gonthier

The Tale of a Brinteaser

Francis Guthrie certainly did it, when he coined his innocent little coloring puzzle in 1852. He managed to embarrass successively his mathematician brother, his brother's professor, Augustus de Morgan, and all of de Morgan's visitors, who couldn't solve it; the Royal Society, who only realized ten years later that Alfred Kempe's 1879 solution was wrong; and the three following generations of mathematicians who couldn't fix it [19].

Even Appel and Haken's 1976 triumph [2] had a hint of defeat: they'd had a computer do the proof for them! Perhaps the mathematical controversy around the proof died down with their book [3] and with the elegant 1995 revision [13] by Robertson, Saunders, Seymour, and Thomas. However something was still amiss: both proofs combined a textual argument, which could reasonably be checked by inspection, with computer code that could not. Worse, the empirical evidence provided by running code several times with the same input is weak, as it is blind to the most common cause of "computer" error: programmer error.

For some thirty years, computer science has been working out a solution to this problem: formal program proofs. The idea is to write code that describes not only *what* the machine should do, but also *why* it should be doing it—a formal proof of correctness. The validity of the proof is an objective mathematical fact that can be checked by a *different* program, whose own validity can be ascertained empirically because it does run on *many* inputs. The main technical difficulty is that formal proofs are very difficult to produce,

Georges Gonthier is a senior researcher at Microsoft Research Cambridge. His email address is gonthier@microsoft.com.

even with a language rich enough to express all mathematics.

In 2000 we tried to produce such a proof for part of code from [13], just to evaluate how the field had progressed. We succeeded, but now a new question emerged: was the statement of the correctness proof (the *specification*) itself correct? The only solution to that conundrum was to formalize the *entire* proof of the Four-Color Theorem, not just its code. This we finally achieved in 2005.

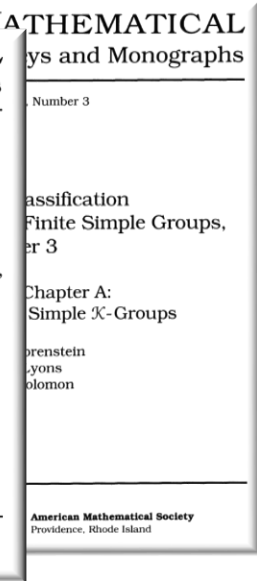
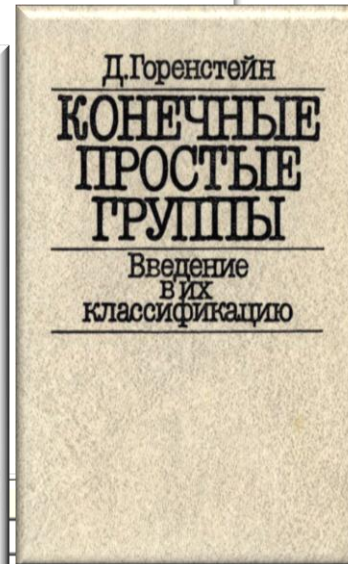
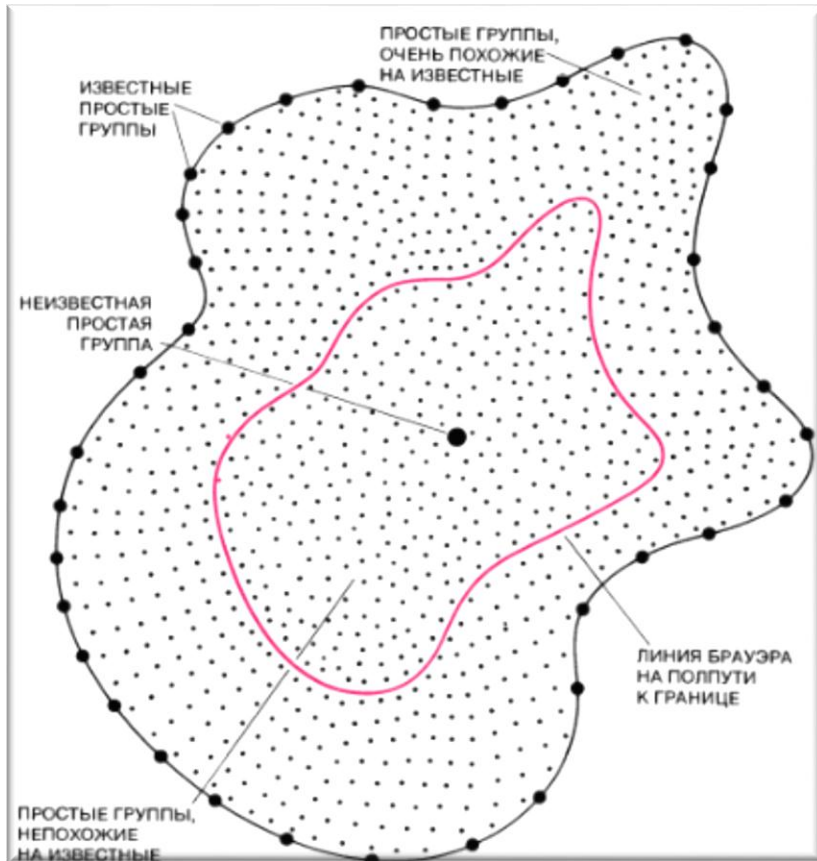
While we tackled this project mainly to explore the capabilities of a modern formal proof system—at first, to benchmark speed—we were pleasantly surprised to uncover new and rather elegant nuggets of mathematics in the process. In hindsight this might have been expected: to produce a formal proof one must make explicit every single logical step of a proof; this both provides new insight in the structure of the proof, and forces one to use this insight to discover every possible symmetry, simplification, and generalization, if only to cope with the sheer amount of imposed detail. This is actually how all of sections "Combinatorial Hypermaps" (p. 1385) and "The Formal Theorem" (p. 1388) came about. Perhaps this is the most promising aspect of formal proof: it is not merely a method to make absolutely sure we have not made a mistake in a proof, but also a tool that shows us and compels us to understand why a proof works.

In this article, the next two sections contain background material, describing the original proof and the Coq formal system we used. The following two sections describe the sometimes new mathematics involved in the formalization. Then the next two sections go into some detail into the two main parts of the formal proof: reducibility and

Большие данные в математике

- Классификация конечных простых групп является одним из основополагающих результатов математики 20-го века.
- Доказательство примерно 10–15 тысяч страниц (примерно 500 статей в научных журналах)
[Д.Горенштейн. *Грандиозная теорема* // Scientific American. Издание на русском языке. – 1986. – № 2. – С. 62–74; R. Solomon. *A brief history of the classification of the finite simple groups* // American Mathematical Society. Bulletin. New Series. – 2001. – V. 38 (3). – P. 315–352.

Большие данные в математике



группа О'Нэна ON	$2^4 \times 3^2 \times 5 \times 11 = 7\,920$
группа Харалды F_2	$2^6 \times 3^3 \times 5 \times 11 = 95\,040$
группа Томпсона F_3	$2^7 \times 2^2 \times 5 \times 7 \times 11 = 442\,570$
группа Фишера F_2 («монстрёнок»)	$2^9 \times 3^2 \times 5^3 \times 7 \times 11 = 44\,352\,000$
группа Фишера-Гривиса F_1 («монстр», «дружественный гигант»)	$2^7 \times 3^6 \times 5^3 \times 7 \times 11 = 898\,128\,000$
спорадическая группа Судзуки Suz	$2^{13} \times 3^7 \times 5^2 \times 7 \times 11 \times 13 \approx 4.48 \times 10^{11}$
	$2^{14} \times 3^3 \times 5^3 \times 7 \times 13 \times 29 \approx 1.46 \times 10^{11}$
	$2^{10} \times 3^3 \times 5^2 \times 7^3 \times 17 \approx 4\,030\,387\,200$
	$2^5 \times 3^7 \times 5^6 \times 7 \times 11 \times 31 \times 37 \times 67 \approx 5.18 \times 10^{16}$
	$2^9 \times 3^4 \times 5 \times 7^3 \times 11 \times 19 \times 31 \approx 4.61 \times 10^{11}$
	$2^{14} \times 3^6 \times 5^6 \times 7 \times 11 \times 19 \approx 2.73 \times 10^{14}$
	$2^{15} \times 3^{10} \times 5^3 \times 7^2 \times 13 \times 19 \times 31 \approx 9.07 \times 10^{16}$
	$2^{41} \times 3^{13} \times 5^6 \times 7^2 \times 11 \times 13 \times 17 \times 19 \times 23 \times 31 \times 47 \approx 4.15 \times 10^{33}$
	$2^{46} \times 3^{20} \times 5^9 \times 7^6 \times 11^2 \times 13^3 \times 17 \times 19 \times 23 \times 29 \times 31 \times 41 \times 47 \times 59 \times 71 \approx 8.08 \times 10^{47}$
	$2^{13} \times 3^7 \times 5^2 \times 7 \times 11 \times 13 \approx 4.48 \times 10^{11}$

Вычислительные машины в математике

- Как пример революции в вычислениях – вычисление числа π
- William Shanks после 15 лет работы в 1873 вычислил 707 цифр числа π , но, как потом выяснилось, только первые 527 оказались верными.
- В 1949 году за 70 часов на компьютере вычислено 2037 знаков числа π : G.W.Reitwiesner. *An ENIAC Determination of π and e to more than 2000 Decimal Places* // *Mathematical Tables and Other Aids to Computation*, Vol. 4, pp. 11 – 15, 1950 (www.jstor.org).
- На самом простом современном компьютере вычисление 100000 знаков числа π займет секунду.

Информационные технологии в математике

- «Ручные» вычисления – типичная практика почти до середины 20 века.
- Вычислительные машины избавили учёных от утомительных арифметических операций.
- Интеллектуальные компьютерные инструменты точно так же должны оставить в прошлом трудоёмкие рутинные (и не только!) операции в математике.

Информационные технологии в математике

- Ожидаем такого же прогресса не только в вычислениях.
- Помимо вычислений и подготовки документов необходимы инструменты интеллектуального поиска, рекомендательные системы формирования близких статей, сервисы терминологического аннотирования, персональные помощники, информационные платформы автоматизации издательской деятельности.

Большие данные и научные коллекции

- Во всем мире число научных статей, согласно подсчету, проведенному Центром исследований в области науки и технологий в Лейденском университете (SBF 2007), значительно возросло. Число научных публикаций в профессиональных журналах во всем мире увеличилось с примерно 686 000 в 1990 году до примерно 1 260 000 в 2006 году, что соответствует увеличению на 84%. Годовой темп роста, рассчитанный на этой основе, составил более 5%. Число научных публикаций растет быстрее, чем мировая экономика

См., например, Binswanger M. *Excellence by Nonsense: The Competition for Publications in Modern Science* // In: Bartling S., Friesike S. (Eds) *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Springer International Publishing, 2014. – P. 49-72. – doi: 10.1007/978-3-319-00026-8_3.

Big Math

Mathematical Intelligencer manuscript No.
(will be inserted by the editor)

Big Math and the One-Brain Barrier A Position Paper and Architecture Proposal

Jacques Carette · William M. Farmer ·
Michael Kohlhase · Florian Rabe

the date of receipt and acceptance should be inserted later

Abstract Over the last decades, a class of important mathematical results have required an ever increasing amount of human effort to carry out. For some, the help of computers is now indispensable. We analyze the implications of this trend towards “big mathematics”, its relation to human cognition, and how machine support for big math can be organized.

The central contribution of this position paper is an information model for “doing mathematics”, which posits that humans very efficiently integrate four aspects: inference, computation, tabulation, and narration around a well-organized core of mathematical knowledge. The challenge for mathematical software systems is that these four aspects need to be integrated as well. We briefly survey the state of the art.

1 Introduction

In the last half decade we have seen mathematics tackle problems that lead to increasingly large developments: proofs, computations, data sets, and document collections. This trend has led to intense discussions about the nature of mathematics, ventilating questions like:

- i) Is a proof that can only be verified with the help of a computer still a mathematical proof?
- ii) Is a mathematical proofscape that exceeds what can be understood in detail by a single expert a legitimate justification of a mathematical result?
- iii) Can a collection of mathematics papers — however big — adequately represent a large body of mathematical knowledge?

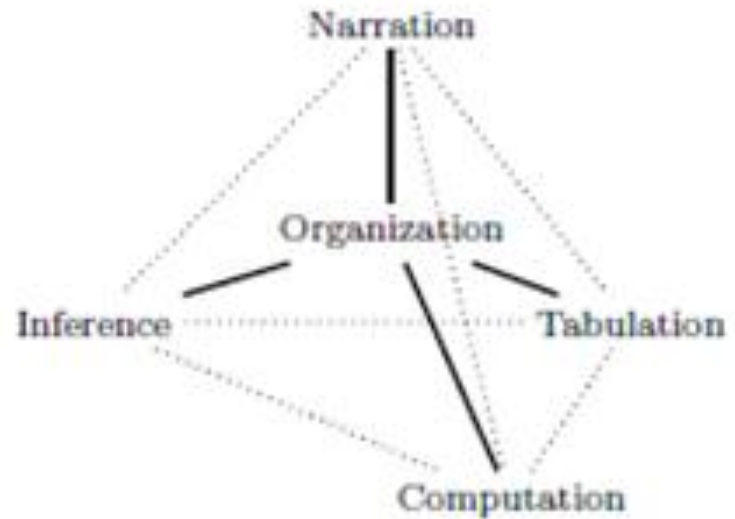
The first question was first raised by Appel and Haken’s proof of the four color conjecture [AHS9] that in 400 pages of regular proof text reduced the problem

arXiv:1904.10405v1 [cs.MS] 23 Apr 2019

Big Math

- В апреле 2019 J. Carette W. M. Farmer, M. Kohlhase, F. Rabe предложили использовать, по аналогии с термином Big Data, термин Big Math для обозначения подхода применения информационных технологий в математике.
- Выделены 5 основных аспекта Big Math: *Inference* (вывод утверждений путем дедукции), *Computation* (алгоритмическое преобразование представлений математических объектов в более легкие для понимания формы), *Tabulation* (создание статических, конкретных данных, относящихся к математическим объектам и структурам, которые можно легко хранить, запрашивать и совместно использовать), *Narration* (приведение результатов в форму, которая может быть усвоена людьми), *Organization* (модульная организация математических знаний)

Big Math



Tetrapod Structure – 5 аспектов системы Big Math Systems.

J. Carette W. M. Farmer, M. Kohlhase, F. Rabe. *Big Math and the One-Brain Barrier. A Position Paper and Architecture Proposal* // arXiv:1904.10405v1 [cs.MS] 23 Apr 2019

Темы по направлению Big Math

- Работы, выполняемые в рамках реализации Программы Центра компетенций Национальной технологической инициативы «Центр хранения и анализа больших данных» МГУ, имеют непосредственное отношение к направлению Big Math.
- Далее выделены некоторые из научных тем, разрабатываемых в настоящее время.

Формализация математического знания

Math Document Mobius Strip

Knowledge	Format	Semantic Model
Text	PDF, LaTeX	-
Metadata	OWL	AKT Portal Ontology
Logical structure	OWL	Mocassin Ontology
Terminology	OWL	OntoMathPro Ontology
Symbolic computation	Coq, Agda	Formalized mathematics



А.М. Елизаров, А.В. Кириллович, Е.К. Липачёв, О.А. Невзорова. *Управление математическими знаниями: онтологические модели и цифровые технологии* // CEUR Workshop Proceedings. – 2016. – Vol. 1752. – P. 44-50. URL: <http://ceur-ws.org/Vol-1752/paper08.pdf>.

Онтологии как формализмы

- Представление и обмен знаниями в любой предметной области основан на ее концептуализации.
- В процессе коммуникации (как между людьми, так и между машинами) используется язык со словарем, содержащим набор терминов для обозначения элементов концептуализации. Для осуществления успешной коммуникации требуется, чтобы все ее участники, во-первых, разделяли общую концептуализацию и, во-вторых, использовали общий словарь. Средством решения этой задачи являются онтологии.

Онтология

- Онтология определяет основные понятия некоторой предметной области и отношения между ними. Основными компонентами онтологии являются: **классы, отношения и аксиомы**.
- Согласно определению Т. Грубера [Gruber T.R. *A Translation Approach to Portable Ontologies* // Knowledge Acquisition. –1993, No 5 (2). – pp. 199–220], онтология – это явная и формальная спецификация разделяемой концептуализации [Studer R., Benjamins R., Fensel D. *Knowledge Engineering: Principles and Methods* // Data & Knowledge Engineering. – 1998. –V. 25, No 1–2. – pp. 161–198].

Онтологии предметных областей науки

- Концепция Семантического Веба предполагает семантическое структурирование пространства интернет-данных для его использования программными агентами, а основными задачами стали унификация (совместимость) и связывание данных из разных источников.
- Наиболее значимым в отношении применения принципов Linked Data является проект LOD. Главное его преимущество – в стандартизованном подходе к структурированию и хранению интегрированных данных, которые загружаются и представляются в виде RDF, т. е. триплетов вида «субъект–предикат–объект».
- Важным направлением области Семантического Веба стала разработка онтологий предметных областей.

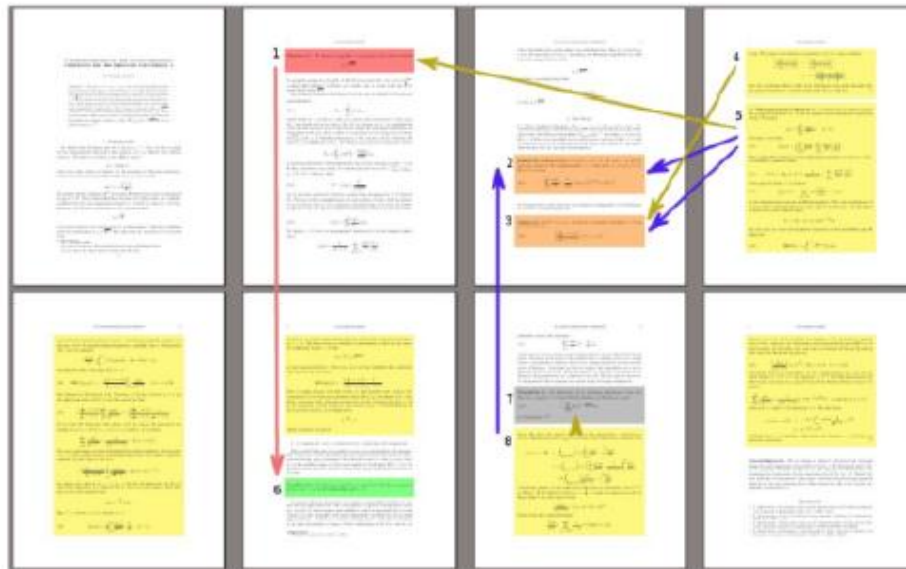
Онтологии математических знаний

- **Mocassin** – онтология логической структуры математических документов, разработанная для автоматического анализа математических публикаций в формате $L^A T_E X$.
- Онтология профессиональной математики **OntoMath^{PRO}**
- Онтология образовательной математики **OntoMath^{EDU}**

А.М.Елизаров, А.Б.Жижченко, Н.Г.Жильцов, А.В.Кириллович, Е.К.Липачёв. *Онтологии математического знания и рекомендательная система для коллекций физико-математических документов*// Докл. Академии наук. – 2016. – Т. 467, №4. – С. 392–395.

Выделение логической структуры

Logical Structure Extraction



Mocassin ontology

- ▼ ● 'Document Segment'
 - Axiom
 - Claim
 - Conjecture
 - Corollary
 - Definition
 - Equation
 - Example
 - Lemma
- ▼ ● 'Mathematical expression'
 - Variable
 - Proof
 - Proposition
 - Remark
 - Theorem

- ▼ ■ topObjectProperty
 - 'depends on'
 - exemplifies
 - 'followed by'
 - 'has consequence'
 - 'has notation'
 - 'has part'
 - 'has segment'
 - proves
 - 'refers to'

- ▼ ■ topDataProperty
 - 'has LaTeX source'
 - 'has start page number'
 - 'has text'
 - 'has title'

Онтология OntoMath^{PRO}

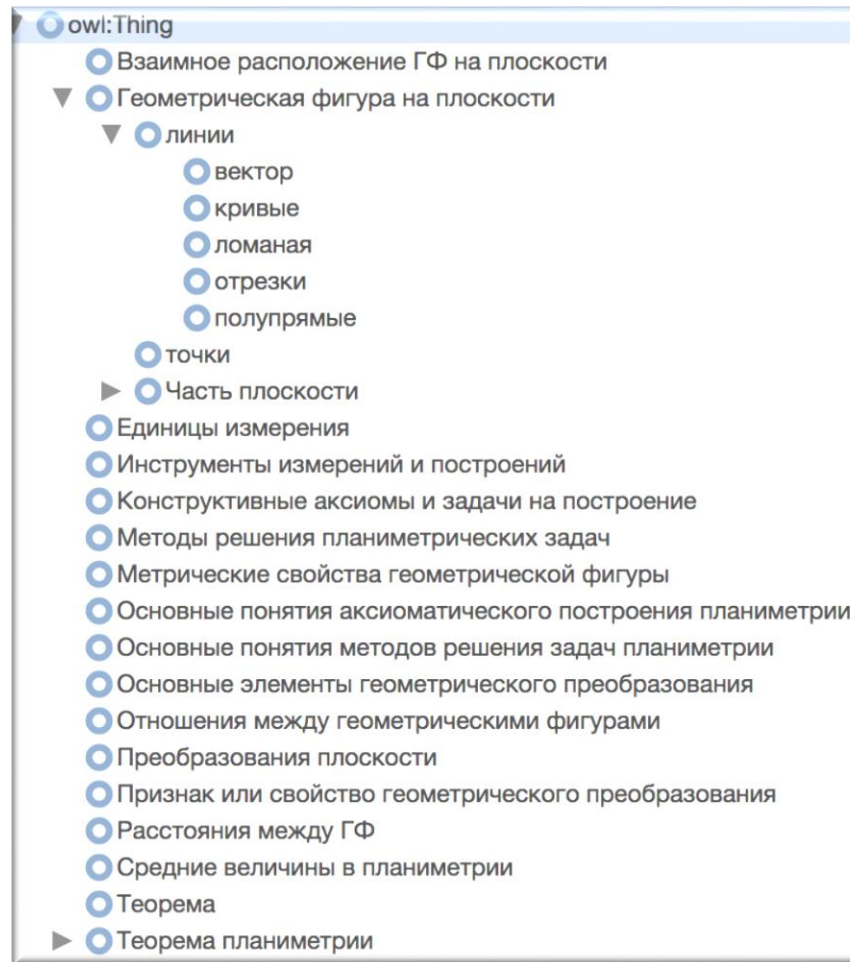
Hierarchy of Objects

- ▼ ● 'Geometric object'
 - ▼ ☉ ' Clifford algebra'
 - 'Even Clifford subalgebra'
 - 'Even component'
 - 'Odd component'
 - ▶ ● Aggregate
 - ▶ ● 'Basic '
 - ▶ ● Curve
 - ▼ ● 'Groupoid '
 - ▼ ● 'Lie algebroid'
 - 'Anchor '
 - ▼ ● 'Topological groupoid '
 - 'Etale groupoid '
 - ▶ ● Point
 - ▼ ● 'Spinor '
 - ' Weyl spinor'
 - 'Dirac spinor'
 - E117
 - 'Isotopic spinor '
 - 'Marjoram spinor'
 - 'Matrix - spinor'
 - Spinor-column
 - ▶ ● Structure
- ▼ ● 'Mathematical knowledge object'
 - ▶ ● Bound
 - ▶ ● Conjecture
 - ☉ Definition
 - ▶ ● 'Element of domain'
 - ☉ 'Element of range'
 - ☉ 'Element of number'
 - ☉ 'Element of point'
 - ☉ 'Element of theory'
 - ☉ 'Element of truth'
 - ☉ 'Element of type'
 - ▶ ● Equation
 - ▶ ● Formula
 - ▶ ● 'Geometric object'
 - ▶ ● Identity
 - ▶ ● Inequality
 - ▶ ● Map
 - ▶ ● Method
 - ▶ ● Number
 - ▶ ● Operation
 - ▶ ● Operator
 - ▶ ● Problem
 - ☉ Proof
 - ▶ ● Property
 - ▶ ● Set
 - ▶ ● Statement
 - ▶ ● Tensor
 - ▶ ● Theorem
 - ▶ ● Transformation
 - ▶ ● 'Value '

Hierarchy of Fields

- ▼ ● 'Field of mathematics'
 - ' Theory of algorithms'
 - ' Theory of differential equations'
 - ' Theory of functions of a complex variable'
 - Algebra
 - Analysis
 - 'Discrete mathematics '
 - 'Functional analysis'
- ▼ ● Geometry
 - ' Analytical geometry'
 - ☉ ' Intrinsic geometry'
 - ' Projective geometry'
 - ▼ ● 'Affine geometry '
 - 'Affine differential geometry '
 - 'Algebraic geometry'
 - ▶ ● 'Differential geometry '
 - 'Fractal geometry'
 - 'Metric geometry '
 - 'Mathematical logic'
 - 'Number theory'
 - 'Numerical analysis'
 - 'Probability theory and mathematical statistics'
 - ▶ ● 'Topology '
- ▶ ● 'Mathematical knowledge object'

Онтология OntoMath^{EDU}



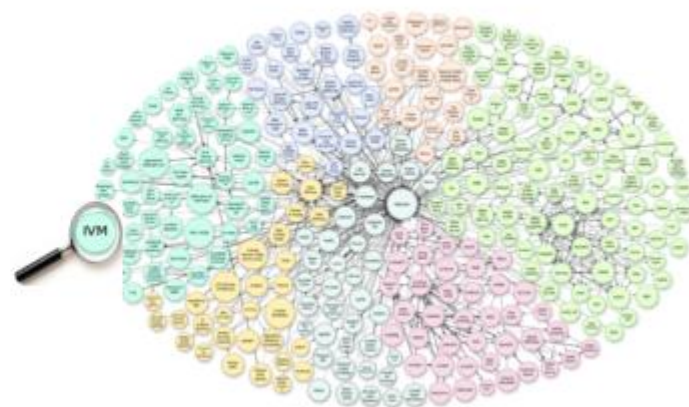
A.Kirillovich, L.Shakirova, M.Falileeva and E.Lipachev. *Towards an Educational Mathematical Ontology* // L. Gómez Chova, et al. (eds). Proc. of the 13th International Technology, Education and Development Conference (INTED2019), Valencia, Spain, March 11th-13th, 2019. IATED, 2019. Pp. 6823-6829.

Платформа семантической публикации

Вход:



Выход:



Коллекция статей в
формате LATEX

RDF-набор

IVM — набор данных LOD, представляющий статьи из журнала «Известия Вузов. Математика» (1997-2007)

Объем: 854 284 триплета, описано 4 190 теорем, 1 015 определений и др.

A. Elizarov, A. Kirillovich, E. Lipachev, O. Nevzorova. *Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management* // Communications in Computer and Information Science, Springer. — 2017. — Vol. 706. — pp 33-46. — doi:10.1007/978-3-319-57135-5_3.

Цифровые математические библиотеки

- World Digital Mathematics Library (WDML)
- Global Digital Mathematics Library (GDML)
- The European Digital Mathematics Library (EuDML)
- MathNet.Ru
- Numerisation de Documents Anciens Mathematiques (NUMDAM), The Czech Digital Mathematics Library (DML-CZ), GDZ–Gottingen Digitization Centre
- Lobachevskii Digital Mathematical Library

A.M.Elizarov, E.K.Lipachev, D.S.Zuev. *Digital mathematical libraries: Overview of implementations and content management services* // CEUR Workshop Proceedings. – 2017. – Vol.2022. – P.317–325.

Lobachevskii DML

<https://lobachevskii-dml.ru/ontomath>

Lobachevskii Digital Mathematics Library

[Home](#) / [OntoMath](#)

OntoMath

OntoMath Digital Ecosystem

OntoMath is a digital ecosystem of ontologies, textual analytics tools, and applications for mathematical knowledge. This system consists of the following components:

- Mocassin, an ontology of structural elements of mathematical scholarly papers;
- OntoMath^{PRO}, an ontology of mathematical knowledge concepts;
- Semantic publishing platform;
- Semantic formula search service;
- Recommender system.



Fig. 1. OntoMath ecosystem architecture

Briefly we describe these basic elements of the architecture of OntoMath digital ecosystem.

The core component of the OntoMath ecosystem is its semantic publishing platform. It builds an LOD representation for a collection of mathematical articles in LaTeX. The generated mathematical dataset includes metadata, terminology, and mathematical formulas. Article metadata, the logical structure of documents in terms of AKT Portal, Mocassin and OntoMath^{PRO} ontologies respectively. Mocassin is an ontology of structural elements of mathematical papers. OntoMath^{PRO} ontologies are parts of OntoMath ecosystem but SALT is an external ontology of the rhetorical structure of mathematical papers. SALT is an external ontology of the rhetorical structure of mathematical papers. SALT is an external ontology of the rhetorical structure of mathematical papers. SALT is an external ontology of the rhetorical structure of mathematical papers.

As any digital ecosystem, OntoMath has components that are used for socio-technical applications and the semantic publishing platform. They can be used by mathematicians and other researchers.

Semantic Publishing Platform

As was mentioned above, the semantic publishing platform which constitutes the core of the OntoMath ecosystem makes an LOD representation for a given sample of mathematical articles in LaTeX. Its main features are:

- Indexing mathematical articles in LaTeX-format as LOD-compatible RDF-data;
- Extracting articles' metadata in terms of AKT Portal Ontology;
- Mining the document logical structure using our ontology of structural elements of mathematical papers;
- Creating instances of mathematical entities as the concepts of OntoMath^{PRO} ontology.

<https://lobachevskii-dml.ru>

Lobachevskii Digital Mathematics Library



Lobachevskii Digital Mathematics Library

Digital Mathematical Library, built on the principle of managing objects of mathematical knowledge, and not mathematical documents. It is based on the fundamental principle of WDML - the principle of creating a network of mathematical information, which is based on knowledge contained in publications presented in electronic collections.

Lobachevskii DML Collections

Digital collection LJM for 1998-2007 with semantic navigation tools. System of relations with the LJM collection from 2007 to the present. Digital collection of Proceedings of the Mathematical Center named after N.I. Lobachevsky, which also includes separate collections of conferences on mathematics and mechanics. Digital collection of the journal *Izvestiya VUZov. Mathematics*. Digital collection of the journal *Uchenye zapiski Kazanskogo Universiteta*.

[View details »](#)

Lobachevskii DML Services

When designing the digital library Lobachevskii-DML we used the results we obtained earlier on the management of mathematical knowledge, as well as developed methods of structural and semantic analysis of mathematical documents. In the LJM electronic collection for 1998-2007, hosted in Lobachevskii-DML, a formula search based on the MathML search method of documents is implemented.

[View details »](#)

OntoMath Ecosystem

OntoMath is a digital ecosystem of ontologies, textual analytics tools, and applications for mathematical knowledge management. The core component of the OntoMath ecosystem is its semantic publishing platform. It builds an LOD representation for a collection of mathematical articles in LaTeX. The generated mathematical dataset includes metadata, the logical structure of documents, terminology, and mathematical formulas.

[View details »](#)

Рекомендательный сервис в цифровой экосистеме OntoMath



А.М.Елизаров, А.Б.Жижченко, Н.Г.Жильцов, А.В.Кириллович, Е.К.Липачёв. *Онтологии математического знания и рекомендательная система для коллекций физико-математических документов*// Докл. Академии наук. – 2016. – Т. 467, №4. – С. 392–395.

A. Elizarov, A. Kirillovich, E. Lipachev, O. Nevzorova. *Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management* // Communications in Computer and Information Science, Springer. – 2017. – Vol. 706. – pp 33-46. – doi:10.1007/978-3-319-57135-5_3.

Семантический рекомендательный сервис

- Контент электронной коллекции семантически анализируется во внешнем сервисе семантического аннотирования Textocat.
- Результаты аннотирования – аннотации в терминах онтологии предметной области OntoMath^{PRO} (<http://ontomathpro.org/>) – сохраняются в базе знаний сервиса.
- Аналитические модули обрабатывают данные из базы знаний и формируют представление в виде интерактивной карточки публикаций и понятий, выводимых на экран пользователя.

OntoMath Semantic Formula Search

The screenshot displays the 'Semantic Formula Search' interface. The main search results page is titled 'Finding Concepts in Mathematical Formulas' and shows the search term 'Gamma function'. Below the search bar, there are filters for various mathematical concepts, including Axiom, Claim, Conjecture, Corollary, Definition, Equation, Example, Lemma, Proof, Proposition, Remark, Theorem, and Other. The search results are presented in a table with columns for Notation, Formula, and Context. The first result shows the Gamma function notation $\Gamma(\cdot)$ and its formula $\Theta_j(s) = \frac{\prod_{k=1}^{m_j} \Gamma(\beta_k^j + s) \prod_{k=1}^{n_j} \Gamma(1 - \alpha_k^j - s)}{\prod_{k=n_j+1}^{p_j} \Gamma(\alpha_k^j + s) \prod_{k=m_j+1}^{q_j} \Gamma(1 - \beta_k^j - s)}$, with an 'Example' context. The second result shows the same notation and formula, but with an 'Other' context. The third result shows the notation $\Gamma(\cdot)$ and the formula $\Theta_j(s) = \frac{\prod_{k=1}^{m_j} \Gamma(\beta_k^j + s) \prod_{k=1}^{n_j} \Gamma(1 - \alpha_k^j - s)}{\prod_{k=n_j+1}^{p_j} \Gamma(\alpha_k^j + s) \prod_{k=m_j+1}^{q_j} \Gamma(1 - \beta_k^j - s)}$, with an 'Example' context. The fourth result shows the notation $\Gamma(\cdot)$ and the formula $\Theta_j(s) = \frac{\prod_{k=1}^{m_j} \Gamma(\beta_k^j + s) \prod_{k=1}^{n_j} \Gamma(1 - \alpha_k^j - s)}{\prod_{k=n_j+1}^{p_j} \Gamma(\alpha_k^j + s) \prod_{k=m_j+1}^{q_j} \Gamma(1 - \beta_k^j - s)}$, with an 'Other' context. A sidebar on the left contains 'Links' (Semantic Math Search, Slides) and 'About' information. A smaller inset window shows a search for 'Ring' with a list of results and a 'Get instances!' button.

A.Elizarov, A.Kirillovich, E.Lipachev, O.Nevzorova. *Semantic formula search in digital mathematical libraries* // Proc. of the 2nd Russia and Pacific Conf. on Comp. Technology and Applications (RPC 2017). IEEE, 2017. Pp. 39-43. DOI: 10.1109/RPC.2017.8168063.

Фабрика метаданных

- Выделение метаданных из документов
- Улучшение метаданных
- Уточнение метаданных
- Пополнение метаданных
- Нормализация метаданных

Автоматизация издательских процессов

- Стилевая валидация поступающих материалов.
- Рекомендательная система подбора классификаторов.
- Сервисы авторов, включая подготовку списков литературы.
- Терминологическое аннотирование статей.
- Рекомендательная система подбора рецензентов.
- Транслитерация списков литературы.
- Формирование метаданных баз цитирования.

Стилевая валидация

Проверка статьи на техническое соответствие

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

**ТЕОРИЯ ВАРИАЦИОННЫХ ОБРАТНЫХ КРАЕВЫХ ЗАДАЧ
АЭРОГИДРОДИНАМИКИ: СОВРЕМЕННОЕ СОСТОЯНИЕ, ПРИЛОЖЕНИЯ,
ПЕРСПЕКТИВЫ РАЗВИТИЯ**

А.М. Елизаров

Казанский (Приволжский) федеральный университет

amehzarov@gmail.com

Аннотация. Вариационные обратные краевые задачи аэрогидродинамики (ОКЗА) реализуют один из подходов к оптимизации аэродинамических и гидродинамических форм, в частности, они связаны с поиском ответа на вопросы, какую максимальную подъемную силу можно получить на профиле крыла и какова форма профилей, обладающих оптимизированными аэродинамическими характеристиками. В рамках классических моделей механики жидкости и газа в математическом плане эти задачи сводятся к вариационным краевым задачам для аналитических функций.

Представлены новые результаты теории вариационных ОКЗА, в том числе близкие к окончательным, описаны приложения в гидродинамике и теории фильтрации, охарактеризованы перспективы развития.

Работа выполнена при поддержке грантов РФФИ №№ 15-07-05380, 15-47-02343.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ТЕОРИИ ВАРИАЦИОННЫХ ОКЗА

Одним из первых примеров вариационной ОКЗА служит задача максимизации подъемной силы дуги заданной длины и ограниченной кривизны при безотрывном ее обтекании потоком идеальной несжимаемой жидкости (ИНЖ). Ее точное решение получено в [1] – доказано, что экстремалью будет дуга окружности. К названному классу относятся многие обобщения обратных краевых задач аэрогидродинамики, теории фильтрации, а также экстремальные задачи теории струй.

Прописные буквы
Полужирный шрифт
Расположен в начале текста

Имеет вид И.О. Фамилия
Следует за названием

Курсив
Расположен после авторов

Размер шрифта: 9пт
Уникальный вид email-а

Начинается со слова «Аннотация»
Размер шрифта: 9пт

Структурный анализ документов

1
2
3
4
5

GENERALIZED ANISOTROPIC NAVIER-STOKES EQUATIONS

S.N. Antontsev¹, H.B. Oliveira²

1 – *University of Lisbon*
2 – *University of Algarve, Portugal*

antontsevsn@mail.ru

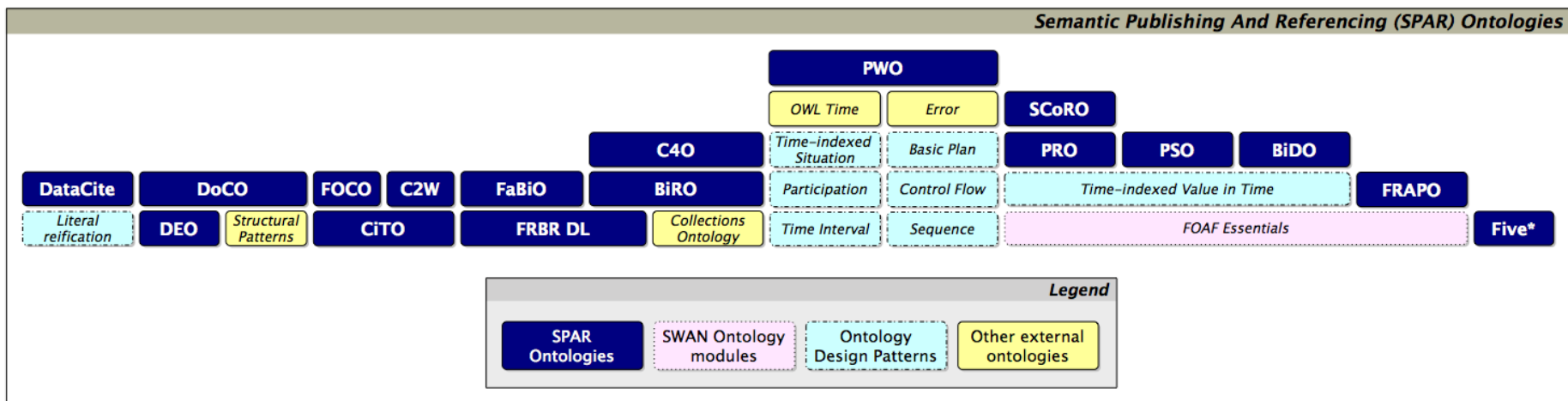
Annotation. In this talk, we consider the evolution problem for the Navier–Stokes equations of non-Newtonian type and with an anisotropic diffusion. It means that the stress tensor is a nonlinear anisotropic tensor function of the strain rate tensor. This set of equations is potentially useful to model fluids that behave as dilatant fluids in several directions and as pseudoplastic in another distinct ones. The existence of weak solutions is proved for the standard initial and boundary value problem. Also the properties of the solutions such as the extinction in a finite time, the exponential time decay and the power time decay are proved. With this respect, we consider the important case of a forces field with different behavior in distinct directions.

INTRODUCTION

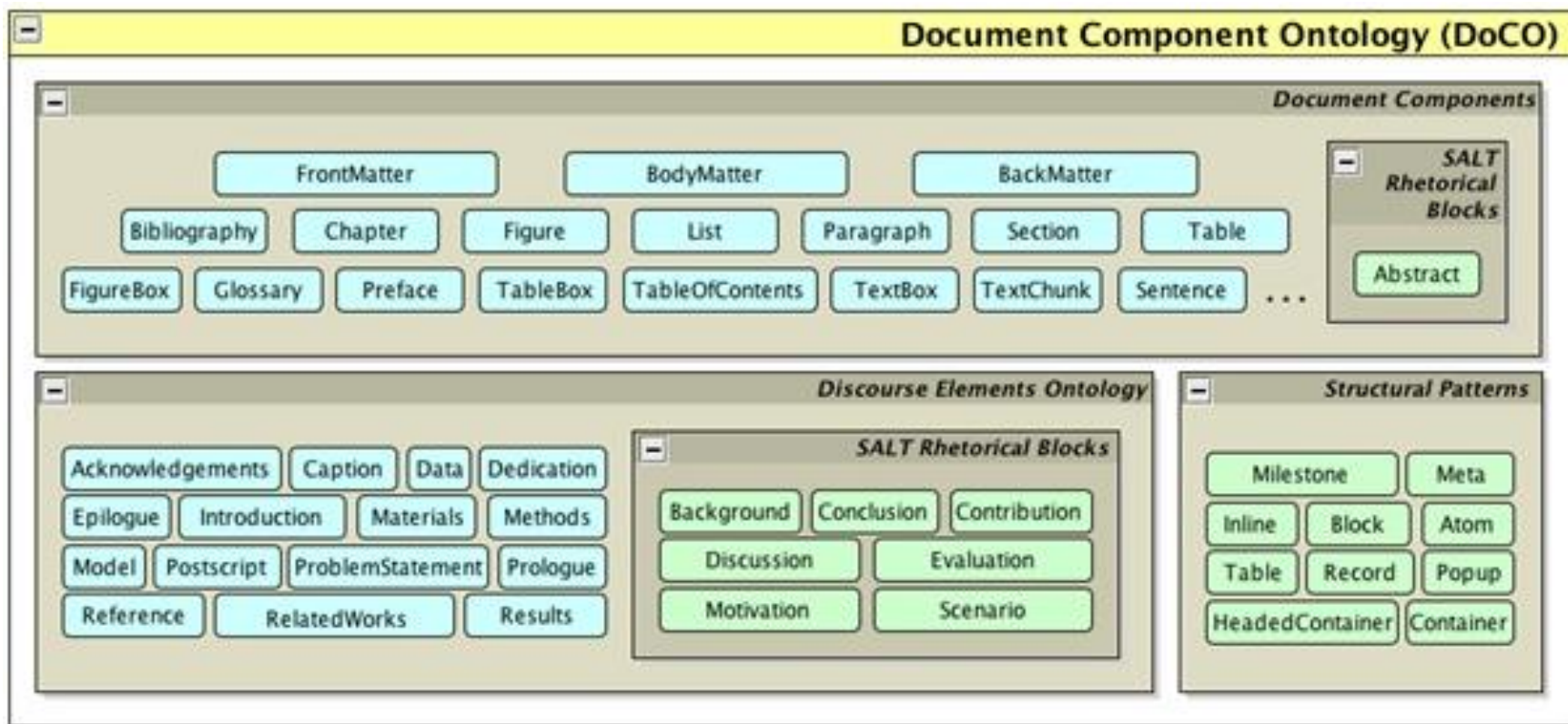
Differential equations with anisotropic diffusion have been studied intensively as witness the large number of works published in recent years (see, e.g. [1] and the references cited therein). In the literature of classical Navier-Stokes equations, anisotropic diffusion is used in the context of geophysical flows with the meaning that the vertical viscosity is distinct from the horizontal one (see e.g. [2, 3]). The resolution of the Navier–Stokes with small initial data in suitable anisotropic spaces is also considered in some literature (see [4] and some of the several works where it is cited). There are also some works in which are studied some anisotropic regularity criteria for the Navier–

А.М.Елизаров, Е.К.Липачёв, Ш.М.Хайдаров. *Автоматизированная система сервисов обработки больших коллекций научных документов // Аналитика и управление данными в областях с интенсивным использованием данных: сборник статей XVIII Междун. конф. DAMDID/RCDL'2016.* – М.: ФИЦ ИУ РАН, 2016. – С.109-115.

Онтологии описания структуры документов



Онтологии описания структуры документов



Структурные элементы

Paper block	Block feature	In ontology terms
Title	Font: Times New Roman, 12 pt, bold, centered etc. Position: at the start of the document	doco:title
Author's list	Font: Times New Roman, 12 pt, centered etc. Position: after title Regex Pattern: authors separated by comma	doco:ListOfAuthors, feof:author
Affiliations	Font: Times New Roman, 12 pt, italic, centered etc. Position: after author's list	pro:relatesToOrganization
E-mail	Font: Times New Roman, 9 pt, bold, centered etc. Position: after affiliations Regex Pattern: Unique address type	fabio:Email
Abstract	Font: Times New Roman, 9 pt, justified etc. Position: after e-mail Regex Pattern: Begins with a specific word: abstract	doco: abstract
References	Position: at the end of the document Regex Pattern: Begins with a specific word: References	doco:bibliography, deo:BibliographicReference

A.Elizarov, S.Khaydarov, E.Lipachev. *Scientific documents ontologies for semantic representation of digital libraries* // Second Russia and Pacific Conf. on Computer Technology and Applications (RPC). Vladivostok, Russky Island, Russia 25-29 September, 2017. – pp. 1–5. DOI: 10.1109/RPC.2017.8168064.

Пример автоматизированного создания авторского указателя

XI Всероссийский съезд по фундаментальным проблемам теоретической и прикладной механики, Казань, 20 – 24 августа 2015 года. С. 4408-4428.

АВТОРСКИЙ УКАЗАТЕЛЬ

Antontsev S.N.	19	Алимова Д.Б.	3215
Chang Ta-Peng	3246	Алисейчик А.П.	132
Ferenczi M.A.	2112	Алокова М.Х.	135
Mace B.R.	869	Аль С.Х.	138
Narayanan T.	2112	Альянов А.В.	141
Oliveira H.V.	19	Алюшин Ю.А.	143
Абдрахманов С.А.	21	Алёхин В.В.	104
Абдрашитов А.А.	24	Амбарцумян Д.С.	687
Абдубакова Л.В.	376	Амелькин Н.И.	145
Абдукаримов А.	28	Аменицкий А.В.	1624
Абдулин А.Я.	1311	Амосов А.С.	4081
Абдулин И.М.	4036	Ананьевский И.М.	147
Абдуллаева М.	172	Ананьевский М.С.	150
Абдуллин А.И.	31	Анджикович И.Е.	563
Абдураимов У.К.	3261	Андреев А.В.	153
Абдуракипов С.С.	34, 2661	Андреев А.Е.	927
Абдыжапар А.	21	Андреев А.С.	156
Абдюшев А.А.	37	Андреев П.С.	1939
Абзалилов Д.Ф.	39	Андронов П.Р.	159
Абрамова К.А.	3626	Андрущенко В.А.	162
Абраров Д.Л.	43	Анисимова И.В.	164
Абрашкин А.А.	46	Анискин В.М.	166
Абрашкин В.И.	49	Анкилов А.В.	726
Абросимов Н.А.	53	Аннакулова Г.К.	169, 172, 3215
Абзязров К.М.	56	Аннин Б.Л.	104, 176

Заключение

- Представлены результаты проекта «*Разработка технологий управления математическими знаниями на основе цифровой математической библиотеки Lobachevskii-DML*», выполняемого в рамках реализации Программы Центра компетенций Национальной технологической инициативы «*Центр хранения и анализа больших данных*», поддерживаемого Министерством науки и высшего образования Российской Федерации по Договору МГУ им. М.В. Ломоносова с Фондом поддержки проектов Национальной технологической инициативы от 11.12.2018 № 13/1251/2018.

Спасибо за внимание!