

# Модель исследовательской инфраструктуры математического знания

А. М. Елизаров<sup>1,2</sup> [0000-0003-2546-6897], Е. К. Липачёв<sup>1</sup> [0000-0001-7789-2332],  
О. А. Невзорова<sup>3</sup> [0000-0001-8116-9446]

<sup>1</sup>*Институт информационных технологий и интеллектуальных систем,  
Казанский федеральный университет*

<sup>2</sup>*Научный исследовательский центр «Курчатовский институт»*

<sup>3</sup>*Институт вычислительной математики и информационных технологий,  
Казанский федеральный университет*

**Аннотация.** Представлена модель инфраструктуры математического знания в виде системы взаимосвязанных математических артефактов, реализованных в виде таких специализированных онтологий, как базовая онтология профессиональных математических знаний, онтология теорем профессиональной математики, онтологии математических задач, методов решения математических задач, алгоритмов и программ, и ряде других. Также в состав модели включены модули автоматизированного построения компонентов инфраструктуры математических знаний и их использование в задачах управления математическим знанием.

**Ключевые слова:** инфраструктура знаний, математический артефакт, математическая онтология, представление математического знания, цифровая библиотека, FAIR.

## Model of the Research Infrastructure of Mathematical Knowledge

A. M. Elizarov<sup>1,2</sup>[0000-0003-2546-6897], E. K. Lipachev<sup>1</sup>[0000-0001-7789-2332],  
O. A. Nevzorova<sup>3</sup>[0000-0001-8116-9446]

<sup>1</sup>*Institute of Information Technologies and Intelligent Systems,  
Kazan Federal University*

<sup>2</sup>*National Research Centre “Kurchatov Institute”*

<sup>3</sup>*Institute of Computational Mathematics and Information Technology, Kazan  
Federal University*

**Abstract.** The model of the infrastructure of mathematical knowledge is presented as a system of interconnected mathematical artifacts implemented in the form of such specialized ontologies as the basic ontology of professional mathematical knowledge, the ontology of theorems of professional mathematics, the ontology of mathematical problems, methods for solving mathematical problems, algorithms and programs, and a number of others. The model also includes modules for the automated construction

of components of the infrastructure of mathematical knowledge and their use in the tasks of managing mathematical knowledge.

**Keywords:** knowledge infrastructure, mathematical artifact, mathematical ontology, representation of mathematical knowledge, digital library, FAIR.

## 1. Введение

Единое пространство научных знаний представляет структурированную интегрированную информационную среду, отражающую достижения различных областей науки [1, 2]. Математическое пространство знаний представляет часть Единого пространства научных знаний, отображающее достижения в математике за весь период развития [3–5].

При формировании цифрового пространства научных знаний необходимо предложить и реализовать методы решения проблемы доступности данных и результатов исследований в научных областях, использующих математические методы.

Данные, используемые при проведении исследований, а также полученные в результате исследований, в настоящее время признаются важной составляющей научного процесса и должны быть проверяемы и доступны для повторного использования. Математические данные имеют специфику, которую необходимо учитывать при разработке программных инструментов. Также математические модели, алгоритмы и программы, как неотъемлемой частью математических исследований, должны быть доступны для обмена, цитирования. Кроме того, исследовательские данные, модели и программы должны быть документированы таким образом, чтобы имелась возможность повторить или воспроизвести представленный результат.

Указанные проблемы могут быть решены путем создания исследовательских инфраструктур. Концепция исследовательских инфраструктур предложена в [6]. Исследовательская инфраструктура представляет систему, объединяющую исследовательские данные, методы и инструменты, а также предоставляющая программные инструменты, позволяющие находить, получать доступ и использовать исследовательские артефакты. Исследовательские математические артефакты являются объектами хранения и состоят из онтологий, таксономий, схем метаданных, датасетов, а также наборов абстрактных сущностей (формулы, теоремы, доказательства и т.д.), содержащиеся в математическом контенте.

В настоящее время общепринятой практикой проектирования исследовательских инфраструктур является использование руководящих документов European Strategy Forum on Research Infrastructures (ESFRI, <https://www.esfri.eu/>). В этих документах предложено следовать принципам FAIR. Общепринятой практикой документирования исследовательских объектов является использование принципов FAIR (**F**indability, **A**vailability, **I**nteroperability, **R**euse), согласно которым при организации данных и сервисов

должны обеспечиваться простота их нахождения (findability), их доступность (availability), их совместимость (interoperability) и возможность их повторного использования (reuse) [7].

В качестве примеров разработки исследовательских инфраструктур укажем проекты [8–10].

Отметим проект Mathematical Research Data Initiative (MaRDI), в котором поставлена задача установить на основе принципов FAIR стандарты для сертифицированных данных математических исследований, проектирования подтверждаемых рабочих процессов и предоставление общественные услуг [10].

В статье [11] поднимается проблема применения принципов FAIR к программному обеспечению. Авторы подчеркивают, что даже на уровне принципов понимание исследовательского программного обеспечения как просто еще одного типа исследовательских данных невозможно.

В работе [9] выделены отличия программного обеспечения от данных в контексте принципов FAIR. Основными отличиями являются частые обновления, наличие связанных объектов, без которых использование программного обеспечения будет невозможно, зависимость от платформы, сложная структура программного обеспечения и, как следствие, необходимость в сложных метаданных для ее описания.

В настоящей работе представлена модель инфраструктуры математического знания в виде системы взаимосвязанных математических артефактов.

## 2. Модель исследовательской инфраструктуры математического знания

На рис. 1 приведена пространственная модель исследовательской инфраструктуры математического знания. Горизонтальные слои представляют компоненты инфраструктуры, а вертикальные – тематические области.



Рис. 1. Модель исследовательской инфраструктуры математического знания

## 2.1. Математические артефакты

Создание цифровой исследовательской инфраструктуры знаний в области профессиональной математики предполагает формирование исследовательских математических артефактов [5, 10, 12–14].

Исследовательские математические артефакты являются объектами хранения и состоят из онтологий, таксономий, графов знаний, схем метаданных, датасетов, а также наборов абстрактных сущностей (формулы, теоремы, доказательства и т.д.), содержащиеся в математическом контенте [5, 15–18].

В целях реализации исследовательской инфраструктуры математического знания спроектированы и разрабатываются следующие математические артефакты.

Математический артефакт в виде новой версии онтологии профессиональной математики  $\text{OntoMath}^{\text{PRO}}$ , основанной на формализме дескрипционной логики, содержащей средства представления математических утверждений, в том числе утверждений, включающих функциональные термы, логические связи и кванторы.

Математические артефакты в виде математических теорем и утверждений, представленных в форме онтологии теорем.

Математический артефакт в виде графа математических формул.

Математический артефакт в виде онтологии математических задач.

Математические артефакты в виде методов решения математических задач, алгоритмов и программ, представленных в форме онтологии методов решения математических задач, алгоритмов и программ.

Математические артефакты в виде датасетов тестовых задач для верификации алгоритмов, созданные в соответствии с принципами FAIR.

Математический артефакт в виде графа знаний для представления организационной структуры математического пространства, включающего описание научных групп, персоналий, исследовательской тематики, представленных в статьях математических журналов.

Математический артефакт в виде онтологической модели представления математического знания как системы взаимосвязанных специализированных онтологий: онтология профессионального математического знания  $\text{OntoMath}^{\text{PRO}}$ , онтология теорем, онтология математических задач, онтология алгоритмов и методов решения математических задач.

## 2.2 . Программные инструменты исследовательской инфраструктуры

Перечислим основные программные инструменты исследовательской инфраструктуры математического знания:

инструменты извлечения именованных сущностей из контента математических документов на основе лингвистических ресурсов и больших языковых моделей (Large Language Models, LLM);

инструменты поддержки ряда задач издательского цикла математического журнала, таких как подбор для статей предметных классификаторов MSC и УДК,

ключевых слов, уточнения и дополнения метаданных, будет выполнена на основе внешних баз знаний (Wikidata, DBpedia, ROR);

рекомендательные системы, формирующие списки тематически близких документов, близких по коду программ, подбора и назначения кода предметных классификаторов для математического документа;

инструменты поддержки процесса научного рецензирования математических документов и программного кода, обеспечивающие автоматизированный подбор рецензентов и другие операции;

### **2.3. Данные и ресурсы исследовательской инфраструктуры**

Перечислим ряд наиболее важных наборов данных исследовательской инфраструктуры математического знания:

таксономия Mathematics Subject Classification 2020 (MSC2020);

Универсальная десятичная классификация (УДК) разделов математики и информатики;

Международный научно-образовательный сайт EqWorld, который содержит обширную информацию о различных классах обыкновенных дифференциальных уравнений, дифференциальных уравнений с частными производными, интегральных уравнений, функциональных уравнений и других математических уравнений;

данные, извлекаемые из статей математических журналов (таблицы, ссылки на датасеты и др.).

### **Заключение**

Представлена модель цифровой исследовательской инфраструктуры знаний в области профессиональной математики, интегрирующая модели представления различных источников математических знаний и учитывающая принципы FAIR. Интегрируемыми моделями являются математические артефакты, включающие онтологии математического знания, датасеты тестовых задач для верификации алгоритмов, граф знаний для представления организационной структуры математического пространства, граф формул.

### **Литература**

1. *Каленов Н.Е., Сотников А.Н.* О структуре онтологии Единого цифрового пространства научных знаний // Научный сервис в сети Интернет: труды XXIV Всероссийской научной конференции (19–22 сентября 2022 г., онлайн). – М.: ИПМ им. М.В.Келдыша, 2022. – С. 203–221. <https://doi.org/10.20948/abrau-2022-23>. <https://keldysh.ru/abrau/2022/theses/23.pdf>

2. *Ataeva O., Kalenov N., Serebryakov V., Sotnikov A.* Informational Infrastructure of the Common Digital Space of Scientific Knowledge // CEUR Workshop Proceedings. 2021. V. 2990. P. 1–10.

3. *Elizarov A., Lipachev E.* Digital Libraries and the Common Digital Space of Mathematical Knowledge // CEUR Workshop Proceedings. 2021. V. 2990. P. 25–38. URL: <https://ceur-ws.org/Vol-2990/rpaper3.pdf>
4. *Елизаров А.М., Кириллович А.В., Липачёв Е.К., Невзорова О.А.* Цифровая экосистема OntoMath: взаимодействие семантических сервисов и математических онтологий // Научный сервис в сети Интернет: труды XXIV Всероссийской научной конференции (19–22 сентября 2022 г., онлайн). М.: ИПМ им. М.В. Келдыша, 2022. С. 167–178. <https://doi.org/10.20948/abrau-2022-40>.
5. *Елизаров А.М., Кириллович А.В., Липачёв Е.К., Невзорова О.А.* Цифровая экосистема OntoMath как подход к построению пространства математических знаний // Электронные библиотеки. 2023. Т. 26. № 2. С. 154–202. <https://doi.org/10.26907/1562-5419-2023-26-2-154-202>
6. *Papou P.* European scientific cooperation and research infrastructures: Past tendencies and future prospects // Minerva. 2004. V. 42. No. 1. P. 61–76, 2004. DOI: 10.1023/B:MINE.0000017700.63978.4a
7. *Wilkinson M. D. et al.* The fair guiding principles for scientific data management and stewardship. // Scientific data. 2016. V. 3. No. 1. P. 1–9. <https://doi.org/10.1038/sdata.2016.18>
8. *Fecher B., Kahn R., Sokolovska N., Volker T., and Nebe P.* Making a research infrastructure: conditions and strategies to transform a service into an infrastructure // Science and Public Policy. 2021. V. 48. No. 4. P. 499–507.
9. *Kalinin N. A. and Skvortsov N. A.* Difficulties of FAIR Principles Implementation in Cross-Domain Research Infrastructures // Lobachevskii J. of Math. 2023. V. 44. No. 1. P. 147–156. DOI: 10.1134/S199508022301016X
10. MaRDI: Mathematical Research Data Initiative Proposal. Version 1. 2022. DOI:10.5281/zenodo.6552436. URL: <https://zenodo.org/records/6552436>
11. *Lamprecht A.-L. et al.* Towards FAIR principles for research software // Data Sci. 2020. V. 3. P. 37–59. <https://doi.org/10.3233/DS-190026>
12. *Elizarov A., Kirillovich A., Lipachev E., Nevzorova O.* Digital ecosystem OntoMath: Mathematical knowledge analytics and management. In: Kalinichenko L., Kuznetsov S., Manolopoulos Y. (eds) Data Analytics and Management in Data Intensive Domains. DAMDID/RCDL 2016. Communications in Computer and Information Science, vol 706. Springer, Cham, 2017. P. 33–46. [http://doi.org/10.1007/978-3-319-57135-5\\_3](http://doi.org/10.1007/978-3-319-57135-5_3)
13. *Schembera B. et al.* Ontologies for Models and Algorithms in Applied Mathematics and Related Disciplines. In: Garoufallou E., Sartori F. (eds) Metadata and Semantic Research. MTSR 2023. Communications in Computer and Information Science. Springer, Cham, 2024. V 2048. P. 161–168. [https://doi.org/10.1007/978-3-031-65990-4\\_14](https://doi.org/10.1007/978-3-031-65990-4_14)
14. *Schembera B. et al.* Towards a Knowledge Graph for Models and Algorithms in Applied Mathematics. In: Sfakakis M. et al. (eds) Metadata and Semantic Research. MTSR 2024. Communications in Computer and Information

Science, Springer, Cham. 2025. V. 2331. P. 95–105. [https://doi.org/10.1007/978-3-031-81974-2\\_8](https://doi.org/10.1007/978-3-031-81974-2_8)

15. *Kirillovich A.V., Nevzorova O.A., Lipachev E.K.* OntoMathPRO 2.0 Ontology: Updates of Formal Model // *Lobachevskii J. of Math.* 2022. V. 43. No. 12. P. 3504–3514. <https://doi.org/10.1134/S1995080222150136>.

16. *Елизаров А.М., Кириллович А.В., Липачёв Е.К., Невзорова О.А.* Онтология математического знания OntoMathPRO // Доклады Российской академии наук. Математика, информатика, процессы управления. 2022. Т. 507. № 1. С. 29–35. <https://doi.org/10.31857/S2686954322700011>

17. *Elizarov A., Kirillovich A., Lipachev E., Nevzorova O., Nevzorov V.* Digital OntoMath Ecosystem Tools for Managing and Developing Mathematical Knowledge // *Lecture Notes in Networks and Systems.* 2024. V. 912. P. 110–117. [https://doi.org/10.1007/978-3-031-53488-1\\_13](https://doi.org/10.1007/978-3-031-53488-1_13)

18. *Петров В.В.* Система автоматизации численной оценки сходства Android-приложений // Научный сервис в сети Интернет: труды XXV Всероссийской научной конференции (18-21 сентября 2023 г.). — М.: ИПМ им. М.В.Келдыша, 2023. — С. 283-297. <https://doi.org/10.20948/abrau-2023-33>.

## References

1. *Kalenov N.E., Sotnikov A.N.* On the structure of the Common Digital Space of Scientific knowledge ontology // *Nauchnyj servis v seti Internet: trudy XXIV Vserossijskoj nauchnoj konferencii (19–22 sentyabrya 2022 g., onlajn).* М.: ИПМ им. М.В. Келдыша, 2022. 2022. – P. 203–221. <https://doi.org/10.20948/abrau-2022-23>.

2. *Ataeva O., Kalenov N., Serebryakov V., Sotnikov A.* Informational Infrastructure of the Common Digital Space of Scientific Knowledge // *CEUR Workshop Proceedings.* 2021. V. 2990. P. 1–10.

3. *Elizarov A., Lipachev E.* Digital Libraries and the Common Digital Space of Mathematical Knowledge // *CEUR Workshop Proceedings.* 2021. V. 2990. P. 25–38. URL: <https://ceur-ws.org/Vol-2990/rpaper3.pdf>

4. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A.* OntoMath Digital Ecosystem: Semantic Service and Mathematical Ontologies // *Nauchnyj servis v seti Internet: trudy XXIV Vserossijskoj nauchnoj konferencii (19–22 sentyabrya 2022 g., onlajn).* М.: ИПМ им. М.В. Келдыша, 2022. P. 167–178. <https://doi.org/10.20948/abrau-2022-40>. <https://keldysh.ru/abrau/2022/theses/40.pdf>

5. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A.* Digital Ecosystem OntoMath as an Approach to Building the Space of Mathematical Knowledge // *Russian Digital Libraries Journal.* 2023. V. 26. No 2. P. 154–202. <https://doi.org/10.26907/1562-5419-2023-26-2-154-202>

6. *Rapon P.* European scientific cooperation and research infrastructures: Past tendencies and future prospects // *Minerva.* 2004. V. 42. No. 1. P. 61–76, 2004. DOI: 10.1023/B:MINE.0000017700.63978.4a

7. *Wilkinson M. D. et al.* The fair guiding principles for scientific data management and stewardship. // *Scientific data*. 2016. V. 3. No. 1. P. 1–9. <https://doi.org/10.1038/sdata.2016.18>
8. *Fecher B., Kahn R., Sokolovska N., Volker T., and Nebe P.* Making a research infrastructure: conditions and strategies to transform a service into an infrastructure // *Science and Public Policy*. 2021. V. 48. No. 4. P. 499–507.
9. *Kalinin N. A. and Skvortsov N. A.* Difficulties of FAIR Principles Implementation in Cross-Domain Research Infrastructures // *Lobachevskii J. of Math.* 2023. V. 44. No. 1. P. 147–156. DOI: 10.1134/S199508022301016X
10. MaRDI: Mathematical Research Data Initiative Proposal. Version 1. 2022. DOI:10.5281/zenodo.6552436. URL: <https://zenodo.org/records/6552436>
11. *Lamprecht A.-L. et al.* Towards FAIR principles for research software // *Data Sci.* 2020. V. 3. P. 37–59. <https://doi.org/10.3233/DS-190026>
12. *Elizarov A., Kirillovich A., Lipachev E., Nevzorova O.* Digital ecosystem OntoMath: Mathematical knowledge analytics and management // *Communications in Computer and Information Science*, vol 706. Springer, Cham, 2017. P. 33–46. [http://doi.org/10.1007/978-3-319-57135-5\\_3](http://doi.org/10.1007/978-3-319-57135-5_3)
13. *Schembera B. et al.* Ontologies for Models and Algorithms in Applied Mathematics and Related Disciplines. In: Garoufallou E., Sartori F. (eds) *Metadata and Semantic Research. MTSR 2023. Communications in Computer and Information Science*. Springer, Cham, 2024. V 2048. P. 161–168. [https://doi.org/10.1007/978-3-031-65990-4\\_14](https://doi.org/10.1007/978-3-031-65990-4_14)
14. *Schembera B. et al.* Towards a Knowledge Graph for Models and Algorithms in Applied Mathematics. In: Sfakakis M. et al. (eds) *Metadata and Semantic Research. MTSR 2024. Communications in Computer and Information Science*, Springer, Cham. 2025. V. 2331. P. 95–105. [https://doi.org/10.1007/978-3-031-81974-2\\_8](https://doi.org/10.1007/978-3-031-81974-2_8)
15. *Kirillovich A.V., Nevzorova O.A., Lipachev E.K.* OntoMathPRO 2.0 Ontology: Updates of Formal Model // *Lobachevskii J. of Math.* 2022. V. 43. No. 12. P. 3504–3514. <https://doi.org/10.1134/S1995080222150136>.
16. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., and Nevzorova O.A.* OntoMathPRO: An Ontology of Mathematical Knowledge // *Doklady Mathematics*. 2022. V. 106 (3). P. 29–35. <https://doi.org/10.1134/S1064562422700016>
17. *Elizarov A., Kirillovich A., Lipachev E., Nevzorova O., Nevzorov V.* Digital OntoMath Ecosystem Tools for Managing and Developing Mathematical Knowledge // *Lecture Notes in Networks and Systems*. 2024. V. 912. P. 110–117. [https://doi.org/10.1007/978-3-031-53488-1\\_13](https://doi.org/10.1007/978-3-031-53488-1_13)
18. *Petrov V.V.* Automated system for numerical similarity evaluation of Android applications // *Nauchnyj servis v seti Internet: trudy XXV Vserossijskoj nauchnoj konferencii (18-21 sentyabrya 2023)*. — M.: IPM im. M.V.Keldysha, 2023. — C. 283-297. <https://doi.org/10.20948/abrau-2023-33>.