



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 71 за 2024 г.



ISSN 2071-2898 (Print)  
ISSN 2071-2901 (Online)

**А.В. Ермаков**

Аналитическая статистика  
научных изданий на Scilit (на  
примере публикаций ИПМ  
им. М.В. Келдыша)

Статья доступна по лицензии  
Creative Commons Attribution 4.0 International



**Рекомендуемая форма библиографической ссылки:** Ермаков А.В. Аналитическая статистика научных изданий на Scilit (на примере публикаций ИПМ им. М.В. Келдыша) // Препринты ИПМ им. М.В.Келдыша. 2024. № 71. 17 с. <https://doi.org/10.20948/prepr-2024-71>  
<https://library.keldysh.ru/preprint.asp?id=2024-71>

**Ордена Ленина  
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ  
имени М.В.Келдыша  
Российской академии наук**

**А.В. Ермаков**

**Аналитическая статистика научных  
изданий на Scilit (на примере  
публикаций ИПМ им.М.В.Келдыша)**

**Москва — 2024**

**Ермаков А.В.**

**Аналитическая статистика научных изданий на Scilit (на примере публикаций ИПМ им. М.В. Келдыша)**

В работе рассмотрены вопросы, связанные с представлением информации о публикациях сотрудников ИПМ и научных изданиях Института в информационно-аналитических материалах системы Scilit. На конкретных примерах показано преимущество полного и корректного задания метаданных научных публикаций, а также продемонстрированы проблемы, возникающие при небрежном обращении с библиографической информацией.

**Ключевые слова:** аналитическая статистика, библиографические базы данных, метаданные публикаций, Scilit.

**Alexey Viktorovich Ermakov**

**Analytical statistics of scientific publications on Scilit (using the example of Keldysh Institute publications)**

The paper examines issues related to the presentation of information about publications by KIAM employees and Institute as scientific publishing house in information and analytical materials of the Scilit system. Specific examples show the advantage of complete and correct setting of metadata for scientific publications, as well as the problems that arise when handling bibliographic information carelessly.

**Key words:** analytical statistics, bibliographic databases, publication metadata, Scilit.

## Оглавление

Введение .....	3
1. Scilit – комплексная платформа-агрегатор для научных публикаций.....	3
2. Издательство ИПМ .....	11
3. Препринты ИПМ.....	14
Заключение.....	16
Список литературы.....	17

## **Введение**

Последние 10 лет показатели публикационной активности российских ученых неуклонно растут. С точки зрения оценки показателей результатов научной деятельности, безусловно, важны индекс цитируемости публикаций и индекс Хирша.

Но с точки зрения автора одним из наиболее важных показателей является интерес читателей (количество просмотров). Многие библиографические системы дают возможность автору увидеть различные статистические показатели не только в статике (общее число просмотров), но и в динамике (за день, за месяц, в виде графиков и т.д.).

Работы по исследованию методов оценки web-аналитики ведутся как в России, так и за рубежом. В российском сегменте это прежде всего проект Science Space от eLibrary [1].

Среди зарубежных следует отметить инструменты Crossref [2] – API от Crossref (<https://www.crossref.org/documentation/retrieve-metadata/rest-api/>) и, конечно же, разработку швейцарско-китайского издательства MDPI (Multidisciplinary Digital Publishing Institute) – комплексную платформу-агрегатор для научных публикаций Scilit.

Издательство MDPI, выпускающее порядка 390 журналов, часто находится на передовой инноваций в области цифровой публикации. Одним из примеров является Scilit [3] – база данных, которая ежедневно собирает, индексирует и извлекает новейшие данные и материалы из CrossRef и PubMed, обеспечивает поиск нужной информации среди 167 миллионов публикаций, 19 миллионов ученых со всего мира, 24 533 издательств, 1 миллиона научных организаций.

Платформа Scilit была разработана в 2014 году. Потенциал этой базы данных позволил достаточно быстро охватить все статьи с цифровым идентификатором объекта (DOI) из Crossref и с идентификатором PMID (PubMed). Scilit анализирует метаданные Crossref и сохраняет их в своей информационной базе. Затем они индексируются в поисковой системе для быстрого использования в режиме реального времени.

В своих аналитических отчетах Scilit четко указывает, является ли материал открытым или нет и какие издатели публикуют работы в удобном и доступном читателю формате. Поскольку академический издательский ландшафт смещается в сторону модели открытого доступа, это очень важный индикатор.

## **1. Scilit – комплексная платформа-агрегатор для научных публикаций**

Главное меню Scilit ([www.scilit.net](http://www.scilit.net)) предлагает выбрать формат представления аналитической информации (рис. 1):

- по публикациям;

- по ученым (авторам);
- по издателям (издательствам);
- по источникам (изданиям);
- по организациям.

Следует сразу отметить, что исходные материалы иллюстраций являются англоязычными, и мы использовали автоматический переводчик Google в представленных ниже иллюстрациях.

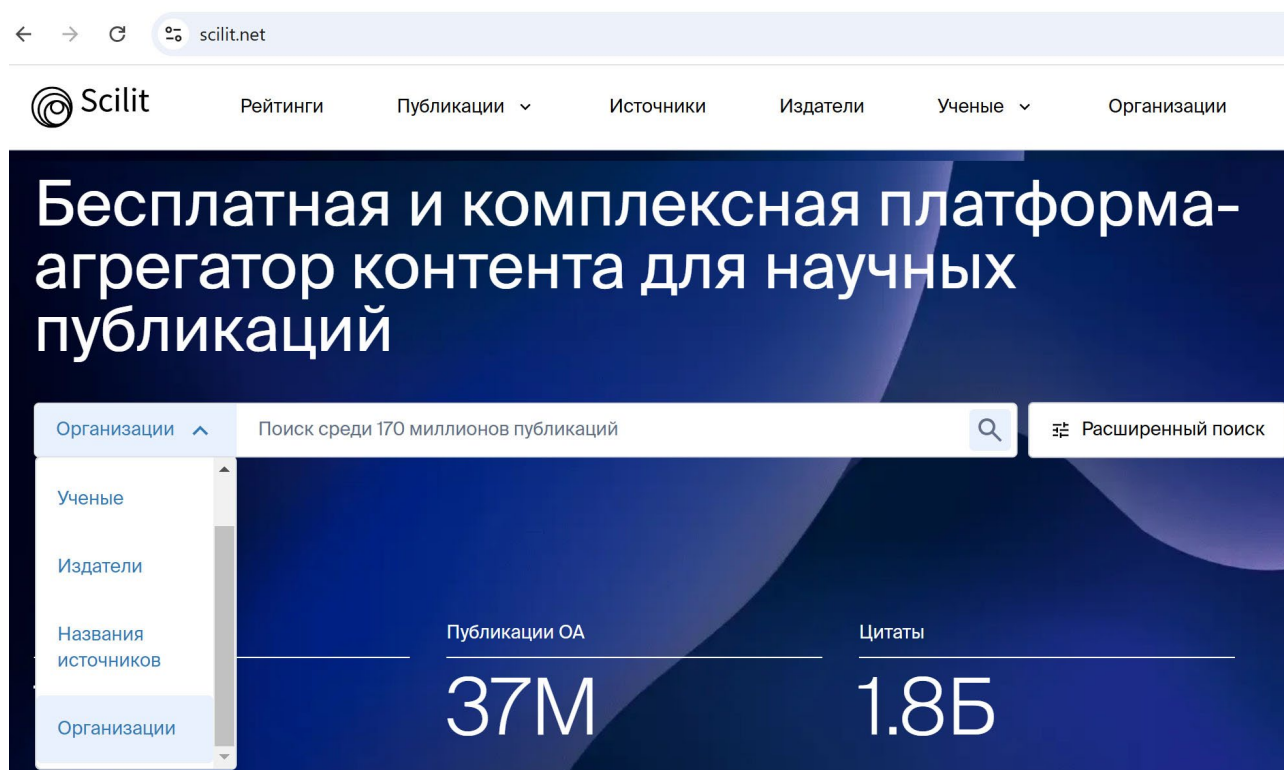


Рис. 1. Главное меню Scilit.

Безусловно, все предлагаемые направления интересны для изучения, но, на наш взгляд, проще и понятней начать рассмотрение с последнего пункта – «Организации».

Выбираем этот пункт и указываем название организации – KIAM (или Keldysh, что еще проще и дает меньше совпадений в поиске).

Здесь следует отметить, что в метаданных параметр «ROR» (Research Organization Registry) появился сравнительно недавно, а до этого авторы не всегда точно указывали в статьях названия организаций, да и издатели далеко не всегда утруждали себя заданием корректной аффилиации авторов. В Препринтах ИПМ этот параметр появился в 2021 году. В некоторых зарубежных и очень малом количестве отечественных журналов – примерно в это же время. А до этого использовалось текстовое название организации ("affiliation": [{"name": "Keldysh Institute of Applied Mathematics"}]). Кто-то указывал полное название, кто-то сокращенное, а многие не указывали вовсе.

В результате значительное количество публикаций наших авторов (сотрудников ИПМ) не попало в статистику, приведенную на рис. 2.



Рис. 2. Анализ публикационной активности сотрудников ИПМ.

Акцент на открытость и доступность публикаций читателям очень важен и является ведущим трендом в издательской области. Справа на рис. 2 представлен график публикаций открытого доступа – публикаций, доступ к которым не связан с лицензионными и коммерческими ограничениями издательств.

Как же получается, что авторы и редакции журналов заинтересованы в попадании своих публикационных материалов в отчеты, статистику, аналитику, но информационно-аналитические системы эти материалы не находят?

Первая причина – это небрежность авторов, опечатки в названии или использование транслитерации вместо корректных англоязычных названий своих организаций. Таким образом аффилиация автора статьи зачастую автоматически переносится в метаданные публикации и отсылается в библиографические базы данных.

Вторая причина – это работа редакции и издательства.

Например, издательство Pleiades Publishing Ltd.

Статья в англоязычной версии журнала Математическое моделирование *Mathematical – Models and Computer Simulations*.

*Kudryashova, T.A., Polyakov, S.V. & Tarasov, N.I. Numerical Analysis of Water Purification Processes in Closed Loop Systems. Math Models Comput Simul 15, 817–831 (2023).*

<https://doi.org/10.1134/S207004822305006X>

Способы получения метаданных из базы Crossref описаны в [4], здесь покажем эту функцию коротко:

```
https://api.crossref.org/works/<DOI>
```

В приведенном примере DOI – 10.1134/S207004822305006X, тогда запрос метаданных указанной статьи в браузере (например, Google) будет выглядеть:

```
https://api.crossref.org/works/10.1134/S207004822305006X
```

Получаем ответ в формате JSON и контекстным поиском находим данные об аффилиации автора:

```
"author":[{"given":"Т.А .","family":"Kudryashova","sequence":
"first","affiliation":[]},
{"given":"S. V.","family":"Polyakov","sequence":
"additional","affiliation":[]},
{"given":"N. I.","family":"Tarasov","sequence":
"additional","affiliation":[]}]
```

Как мы видим, в метаданных этой статьи задана «пустая» аффилиация авторов.

Другой способ исследования метаданных научных публикаций – анализ информации в формате XML.

Формат запроса метаданных указанной статьи в этом случае будет выглядеть так:

```
https://api.crossref.org/works/10.1134/S207004822305006X.xml
```

В полученном ответе находим информацию об авторах (contributors):

```
<contributors>
<person_name contributor_role="author" sequence="first">
<given_name>Т. А.</given_name>
<surname>Kudryashova</surname>
</person_name>
<person_name contributor_role="author" sequence="additional">
<given_name>S. V.</given_name>
<surname>Polyakov</surname>
</person_name>
<person_name contributor_role="author" sequence="additional">
<given_name>N. I.</given_name>
<surname>Tarasov</surname>
</person_name>
</contributors>
```

Но никакой информации о месте работы авторов в представленном фрагменте метаданных нет.

При этом и в электронной версии статьи на сайте Springer, и в pdf-версии статьи указаны правильные аффилиации – Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, Russia.

Другой пример не совсем корректного подхода к указанию аффилиации автора – издательство MDPI.

Регистрируясь на сайте издательства [5], каждый автор заполняет шаблон (рис. 3, 4).

Рис. 3. Шаблон для регистрации автора в издательстве MDPI.

Рис. 4. Поля шаблона, связанные с аффилиацией.

Рассмотрим статью в журнале Aerospace (MDPI):

*Vladislav Orlov, Uliana Monakhova, Mikhail Ovchinnikov and Danil Ivanov. Fuelless On-Orbit Assembly of a Large Space Truss Structure Using Repulsion of the Service Spacecraft by Robotic Manipulators*

<https://doi.org/10.3390/aerospace11080635>

В метаданных статей указанное издательство в поле аффилиации записывает все обязательные поля своего регистрационного шаблона (рис.4) – Affiliation, Address1, Zip Code, City, Country.

```
affiliation": [{"name": "Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, 125047 Moscow, Russia"}]
```

Конечно, это избыточно, но, безусловно, лучше, чем «пустая» аффилиация. Информационно-аналитическая система при обработке этой



информации и читатель при изучении аналитики могут узнать не только организацию автора, но и город, и страну проживания.

Анализ публикации От 2014 к 2024

**Источники**

Публикации в 2014-2024 гг.

Название источника	Публикации
Препринты Института Келдыша	914
Математические модели и компьютерное моделирование	183
Журнал физики: Серия конференций	68
Журнал математики Лобачевского	42
Космические исследования	38
Исследования Солнечной системы	30
Программирование и компьютерное программное обеспечение	29
Вычислительная математика и математическая физика	27
Математическое Моделирование	23
Акта Астронавтика	17

Рис. 5. Источники метаинформации о публикациях сотрудников ИПМ.

На рис. 5 показаны «источники» метаинформации – издания, в которых публиковались статьи сотрудников ИПМ и которые стали источником метаданных научных публикаций для приведенной статистики.

На рис. 6 показана «тематическая направленность» наших публикаций. Но в метаданных тематическое направление исследований не задается.

На основе модели машинного обучения Scilit автоматически классифицирует публикации по различным темам. Модель классификации тем была обучена только на англоязычных публикациях и может привести к «слабым» результатам для публикаций на других языках (если тематическое направление вообще можно определить).

Таким образом, система сама, возможно по названию, ключевым словам или аннотациям, построила таблицу тематических направлений.

## Главные темы

Публикации в 2014-2024 гг.

Предмет	Публикации
Компьютерная инженерия	<u>185</u>
Астрофизика	<u>159</u>
Лазерная физика	<u>80</u>
Физика плазмы	<u>51</u>
Статистическое моделирование	<u>47</u>
Физическое образование	<u>38</u>
Функциональный анализ	<u>35</u>
Статистическая физика	<u>32</u>
Физика атмосферы	<u>30</u>
Численные методы	<u>25</u>

Рис. 6. Тематическая направленность публикаций сотрудников ИПМ.

В целом, это безусловно хорошо и удобно исследователю. Однако на данный момент корректность отнесения той или иной публикации к определенной теме вызывает вопросы.

Это легко проверить: справа указано количество публикаций, отнесенных к данной теме. Это поле является «кликабельным», т.е. можно получить весь список статей.

Например, система ИИ отнесла к Астрофизике статью:

*Галанин М.П., Лукин В.В., Родин А.С.*

*Программная платформа Temetos и ее применение в задачах механики сплошной среды*

<https://doi.org/10.1134/s199047892304004x>

Также вызывает определенный интерес статистика цитирования (рис. 7), которую можно получить за любой диапазон лет с начала наполнения базы данных Scilit (2014 год).

## Анализ цитирования (?)

**Данные по цитированию за все время.**

Всего цитирований	16,310
Всего публикаций	3,090
Среднее количество цитирований	5.28
Цитируемые публикации	1,882
Процитированные публикации %	60,9%
Самоцитирование	2,376
% самоцитирования	14,6%

Рис. 7. Статистика цитирования публикаций сотрудников ИПМ.

«Кликнув» на вопросительный знак «?» в верхней части рисунка, можно получить подробную информацию по каждой позиции данного статистического отчета.

На рис. 8 показаны данные по цитированию в графическом виде.

Все цитаты за 2014-2024 гг.

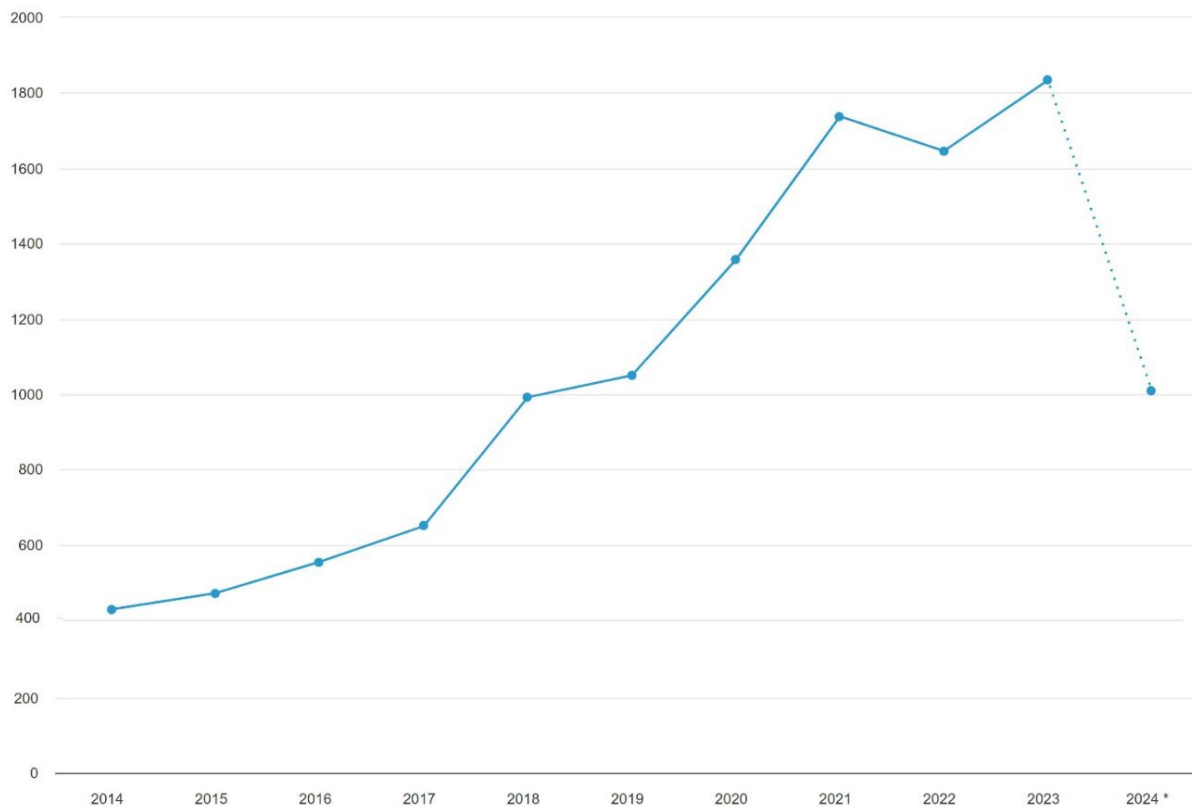


Рис. 8. Графическое представление статистики цитирования публикаций сотрудников ИПМ.

Еще одним интересным, на наш взгляд, показателем являются данные о сотрудничестве – соавторстве с сотрудниками других организаций (рис. 9).



Рис. 9. Данные по научным публикациям, подготовленным в соавторстве с сотрудниками других организаций.

## 2. Издательство ИПМ

Теперь рассмотрим статистику публикаций, которые вышли в издательстве ИПМ: препринты, журналы, монографии, сборники конференций (рис. 10). Для этого необходимо вернуться на главную страницу системы Scilit, выбрать «Издатели» (Publishers) и задать «Keldysh Institute».

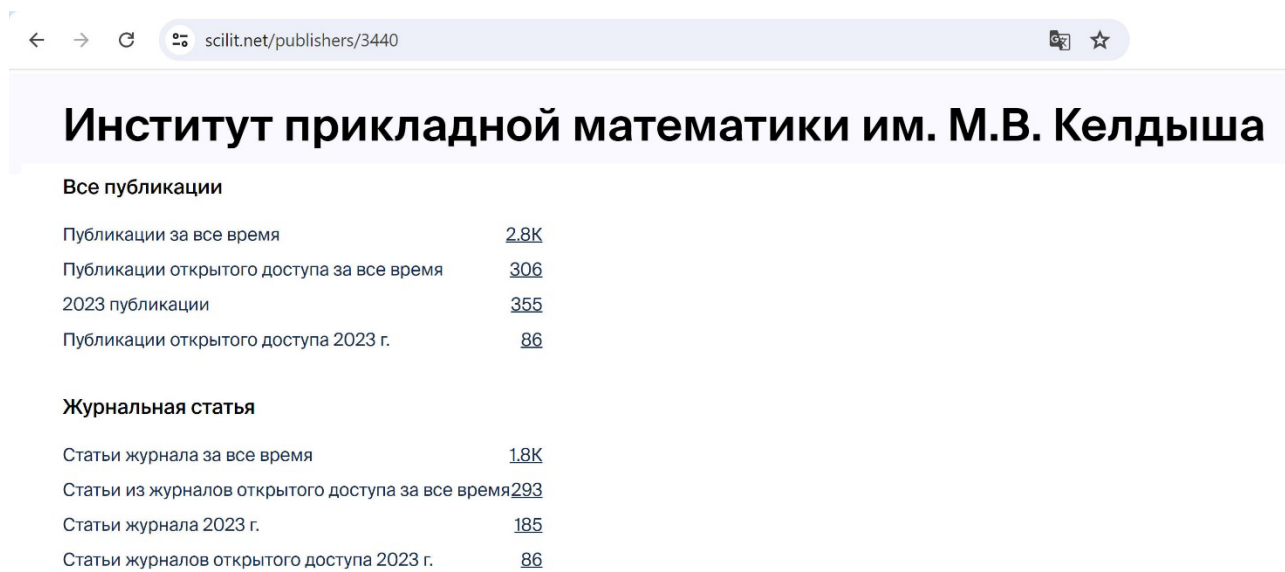


Рис. 10. Статистика по изданиям ИПМ.

Далее на рис. 11 представлены наши издания – источники информации для данного аналитического обзора.

### Источники

Публикации в 2014-2024 гг.

Название источника	Публикации
Препринты Института Келдыша	1,234
Математическое Моделирование	354
Математика Монтиснигри	156
Труды Международной конференции «Проектирование будущего. Проблемы цифровой реальности»	9
Научный сервис в сети интернет.	2

Рис. 11. Издания ИПМ.

Естественно, хотелось бы сразу разобраться в точности и корректности представленной информации. «Кликнув» на значение для Препринтов Института Келдыша (1234), получаем список, в котором последний опубликованный по состоянию на 13.11.2024 препринт – № 66 от 2024 г.

Опубликовано Институт прикладной математики им. М.В. Келдыша в Препринты Института Келдыша № 66 , стр. 1-16

<https://doi.org/10.20948/prepr-2024-66>

Моделирование течения углеводородной жидкости в каталитическом реакторе

Глюс Лябибович Шарипов, Камила Феликсовна Коледина, Виктория Олеговна Подрыга, Сергей Владимирович Поляков, Никита Игоревич Тарасов

На самом деле № 67 был опубликован 11.11.2024, но это приемлемая, на наш взгляд, задержка разработчиков системы, хотя она противоречит их утверждениям о ежедневном сканировании обновлений Crossref.

Для нас такого рода несоответствие вполне ожидаемо: 6-7 лет назад базы данных Crossref и Scilit были значительно меньше. Сейчас они превзошли терабайтные объемы, и обеспечение высокой скорости обработки поступающих обновлений стало значительно сложнее.

Безусловно, интересна тематическая направленность публикаций в изданиях ИПМ (рис. 12).

#### Главные темы

Публикации в 2014-2024 гг.

Предмет	Публикации
Управление рисками и оценка	<u>53</u>
Компьютерная инженерия	<u>45</u>
Социальная справедливость и реформа	<u>44</u>
Математическая Теория	<u>35</u>
Функциональный анализ	<u>24</u>
Астрофизика	<u>23</u>
Графовые алгоритмы и структуры	<u>20</u>
Компьютерное зрение и графика	<u>19</u>
Взаимодействие человека и машины	<u>19</u>
Информационная безопасность	<u>18</u>

Рис. 12. Тематический каталог публикаций в изданиях ИПМ.

Эта тематическая таблица также построена при помощи ИИ, и к ней также возникают вопросы. Но, скорее всего, с развитием модели ИИ эти вопросы будут постепенно сниматься.

Однако здесь необходимо сделать одно важное замечание. Так как журнал «Математическое моделирование» не размещает в метаданных статей англоязычные названия и аннотации, то в аналитических обзорах Scilit эти материалы очень нескоро будут отражены в тематических рубриках.

Безусловно, интересной является статистика по сотрудничеству наших ученых с зарубежными специалистами (рис. 13).

Публикации в 2014-2024 гг.

Страна	Публикации
Россия	<u>1,401</u>
Алжир	<u>23</u>
Беларусь	<u>10</u>
Турция	<u>10</u>
Черногория	<u>9</u>
Иран	<u>7</u>
Таджикистан	<u>6</u>
Украина	<u>6</u>
Армения	<u>3</u>
Азербайджан	<u>3</u>

Рис. 13. Сотрудничество с зарубежными специалистами.

Основной вклад в «географию» авторов наших публикаций делает журнал «Математика Черногории» (Mathematica Montisnigri), издаваемый нашим Институтом совместно с Обществом математиков и физиков Черногории и Кафедрой математики Университета Черногории.

### 3. Препринты ИПМ

Анализ статистики по изданиям Института покажем на примере Препринтов ИПМ. Для этого нужно снова вернуться на главную страницу системы Scilit, выбрать «Названия источников» (Source Titles) и задать «Keldysh Institute Preprints».

На рис. 14 представлена графическая информация за все годы нашего сотрудничества с Crossref (метаданные Препринтов ИПМ стали загружаться в Crossref с 2016 года). Эта наша работа достаточно четко отражена на графике Scilit.

## Анализ публикаций

Публикации в 2014-2024 гг.

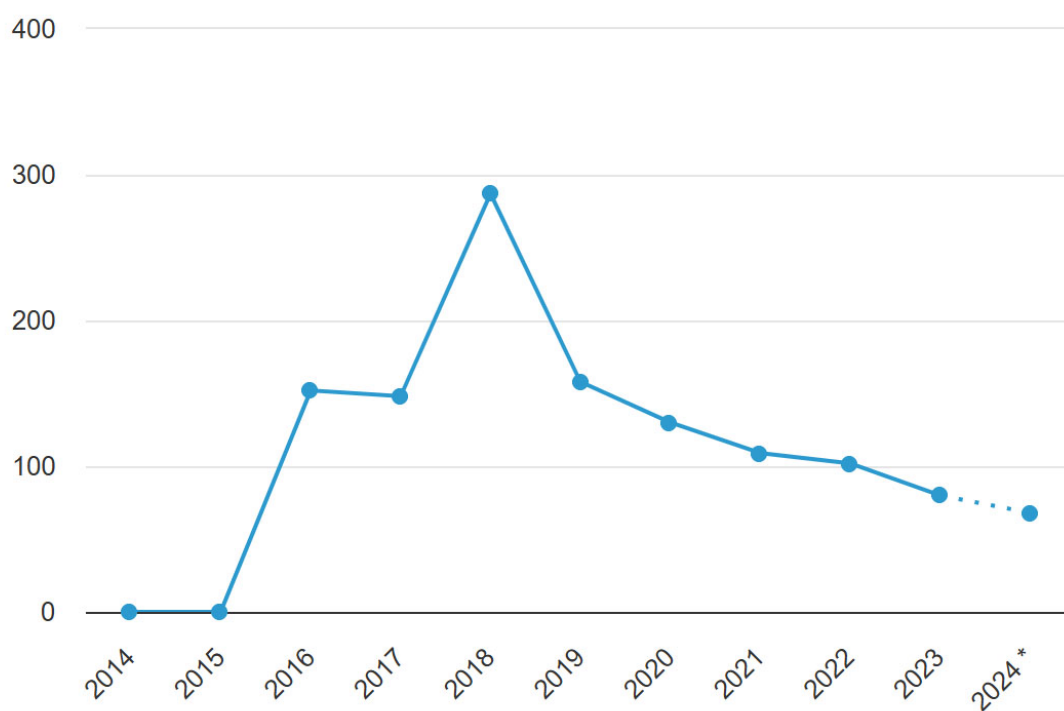


Рис. 14. Статистический анализ по Препринтам ИПМ.

Естественно, у нас были возможность и желание проверить каждую точку на графике. Оказалось – придаться не к чему. Например, препринтов 2023 года 80 – было выпущено 78 русскоязычных и плюс 2 англоязычных – № 26-е (Виктора Дмитриевича Лахно) и № 52-е (Александра Ивановича Аптекарева).

Заинтересовало также международное сотрудничество в Препринтах (рис.15).

### Лучшие страны

Публикации в 2014-2024 гг.

Страна	Публикации
Россия	<u>921</u>
Таджикистан	<u>3</u>
Италия	<u>1</u>
Узбекистан	<u>1</u>

Рис. 15. Международное сотрудничество авторов Препринтов ИПМ.



Интересно, кто из зарубежных коллег «отметился» в Препринтах ИПМ. Начнем с итальянца.

*Published by Keldysh Institute of Applied Mathematics in Keldysh Institute Preprints No. 150-e, p. 1-24*

<https://doi.org/10.20948/prepr-2019-150-e>

*Processes of microporosity formation at initial stage phase transition*

*by Galina Ivanovna Zmievskaya, Tatiana Viktorovna Levchenko, Giuseppe Maino*

В метаданных этого препринта указана аффилиация Giuseppe Maino:

*"affiliation": [{"name": "Bologna University, Scuola di Lettere e Beni Culturali, Italy"}]*

И система совершенно справедливо отнесла его к Италии.

Следующим рассмотрим список соавторов из Узбекистана:

*Published by Keldysh Institute of Applied Mathematics in Keldysh Institute Preprints No. 106, p. 1-12*

<https://doi.org/10.20948/prepr-2019-106>

*Application of the fictitious domains method for the simulation of gas flows at low Mach numbers*

*by Marina Alexandrovna Trapeznikova, Natalia Gennadievna Churbanova, Gairat Umarovich Juraev*

Здесь также безусловно точно Гайрат Умарович Жураев отнесен к Узбекистану. Это произошло наверняка не без помощи ИИ, так как нигде в метаданных явно не указана страна. Ее можно «вычлениить» только из названия Университета.

*"affiliation": [{"name": "National University of Uzbekistan named after Mirzo Ulugbek"}]*

Еще одна позиция в таблице на рис. 15 – Таджикистан, к представителям которого совершенно необъяснимым образом система отнесла соавтора 3 препринтов, сотрудника МГУ им.М.В.Ломоносова Дениса Михайловича Буланова.

## Заключение

В данной работе мы не ставили задачу подробно рассказать обо всех возможностях информационно-аналитической системы Scilit. Нам было важно показать на конкретных примерах некоторые наиболее интересные, на наш взгляд, ее возможности. Показать, какие преимущества дает аккуратное и наиболее полное задание метаданных научных публикаций и какие проблемы возникают при небрежном отношении авторов и/или издателей к этой информации.

## Список литературы

1. Информационно-аналитическая система SCIENCE SPACE.  
[https://elibrary.ru/projects/science\\_space/science\\_space\\_info.asp?](https://elibrary.ru/projects/science_space/science_space_info.asp?)
2. API унифицированных ресурсов CrossRef.  
<https://api.crossref.org/swagger-ui/index.html>
3. Платформа-агрегатор для научных публикаций Scilit.  
<https://www.scilit.net>
4. Ермаков А.В. Некоторые проблемы при регистрации даты публикации в Crossref // Научный сервис в сети Интернет: труды XXIV Всероссийской научной конференции (19-22 сентября 2022 г., онлайн). — М.: ИПМ им. М.В.Келдыша, 2022. — С. 179-182.  
<https://doi.org/10.20948/abrau-2022-21>  
<https://keldysh.ru/abrau/2022/theses/21.pdf>
5. Регистрационная форма автора в издательстве MDPI  
<https://susy.mdpi.com/user/edit>