

УДК 004

ПОВЫШЕНИЕ КАЧЕСТВА МЕТАДААННЫХ НАУЧНЫХ ПУБЛИКАЦИЙ С ПОМОЩЬЮ ОТЧЕТОВ CROSSREF

А. В. Ермаков^[0000-0002-6054-0813]

Институт прикладной математики им. М.В. Келдыша Российской академии наук,
Миусская пл., 4, Москва, 125047

Ermakov@Keldysh.ru

Аннотация

Рассмотрены вопросы, связанные с повышением качества метаданных научных публикаций, размещаемых в библиографической базе данных Crossref. Всю информацию, содержащуюся в метаданных, полученных от издателей научных публикаций, Crossref анализирует и отображает в различных отчетах. Отчеты дают издателям представление о полноте и корректности представленных библиографических данных. Качество метаданных прямо или косвенно влияет на количество просмотров и ссылок на публикацию, соответственно, на рейтинги научных изданий, авторов и организаций.

Ключевые слова: метаданные публикаций, отчеты Crossref, цитирование, рейтинги научных изданий.

ВВЕДЕНИЕ

Подготовка научной публикации неразрывно связана с тематикой и направлением исследований, в которых работают авторы. Как правило, публикации являются продолжением работ научного коллектива в данном направлении и опираются на предыдущие результаты исследований, выполненных ранее либо самими авторами, либо их научными руководителями или коллегами. Эта опора в публикациях на предшествующие результаты научных исследований сопровождается цитированием соответствующих научных материалов. Корректность цитирования очень важна не только для авторов новой научной статьи, но и для авторов цитируемых статей, а также для их научных изданий, так как прямо или косвенно влияет на количество просмотров и ссылок на публикацию, соответственно, на рейтинги научных изданий, авторов и организаций.

Метаданные публикаций необходимы, чтобы сделать библиографические данные о самой статье, авторах и т. п. идентифицируемыми, открываемыми и доступными для других.

Всем публикуемым научным материалам присваивается цифровой идентификатор DOI (Digital Object Identifier), обеспечивающий ссылку (URL) на постоянное местонахождение объекта или информацию о нём (метаданные) в интернете.

Подавляющее число авторов имеют ORCID – это уникальный код (ID), который автор научных работ получает для идентификации полученных им результатов и написанных трудов. Главная задача ORCID – однозначное определение исследователя во всех библиографических базах данных.

Ассоциация Crossref [1], членом которой с 2016 г. является Институт прикладной математики (ИПМ) им. М.В. Келдыша, поддерживает совместную всемирную службу взаимной цитируемости, функционирующую как своеобразный шлюз между электронными платформами издателей. Эта служба не хранит полные тексты научных публикаций, но заносит в свою базу данных информацию о связи публикаций с помощью технологии DOI [2], а также метаданные опубликованных научных материалов.

Разрабатываемые инструменты Crossref (и некоторых других организаций, таких как Google Scholar, Scopus и Web of Science, которые используют различные источники для своих данных цитирования) облегчают как автору публикации, так и читателям поиск, цитирование, оценку, повторное использование результатов научных исследований. Кроме того, статистика «считывания» (переходов на публикацию), а также обратных цитирований [3] (ссылок на публикации, ссылающиеся на данную статью) в той или иной мере оказывают влияние на рейтинг научных сотрудников и всего Института в целом. Поэтому ИПМ и как издателю научных материалов, и как научно-исследовательскому центру важно отслеживать корректность метаданных выпускаемых научных публикаций.

Для проверки качества метаданных Crossref разработал достаточно большой набор инструментов, которые помогают издателю оценить и улучшить свои метаданные.

В данной работе рассмотрены отчеты, к которым любой издатель или заинтересованный автор может получить доступ на сайте Crossref. Анализ инфор-

мации, представленной в этих отчетах, позволяет качественно оценить полноту отгружаемого контента, связанного с издаваемыми научными материалами, и наметить пути улучшения инфраструктуры научных публикаций и инструментов взаимодействия с Crossref.

1. СПИСОК ИЗДАНИЙ С ДОСТУПНЫМИ ДЛЯ ПРОСМОТРА МЕТАДААННЫМИ

Подробно расскажем об отчетах Crossref на примере некоторых научных изданий ИПМ и Казанского федерального университета (КФУ). На примере изданий ИПМ покажем, как издательство может с помощью отчетов Crossref находить ошибки и/или неточности в метаданных публикаций и вносить необходимые исправления. На примере изданий КФУ будет показано, как заинтересованный автор может обнаружить неточности или неполноту метаданных интересующего его издания. Далее автор может обратиться к издателю, указав на неточности, и попросить их исправить.

Любое научное издательство может таким же образом изучить отчеты Crossref, фокусируя внимание на своих научных изданиях.

Всем организациям, сотрудничающим с Crossref, предоставляется возможность не только увидеть информацию о своих изданиях, но и внести необходимые корректировки.

Информация представляется в виде трех списков, упорядоченных по алфавиту:

- Журналы

<https://www.crossref.org/06members/51depositor.html>

- Сборники конференций

<https://www.crossref.org/06members/51depositorCP.html>

- Книги

<https://www.crossref.org/06members/51depositorB.html>

На рис. 1 представлен фрагмент алфавитного списка мировых издателей журналов, среди которых ИПМ (Keldysh Institute of Applied Mathematics) представлен тремя журналами:

- Mathematica Montisnigri (Математика Черногории);
- Mathematischesкое моделирование (Математическое моделирование);
- Keldysh Institute Preprints (Препринты ИПМ).

The screenshot shows a web browser window with the URL <https://www.crossref.org/06members/51depositor.html>. The page displays a list of publishers and their associated journals. The table below summarizes the data visible in the screenshot.

Journal	# DOIs	Last Crawl Date
Kelam Arastirmalari Dergisi (Journal of Kalam Studies)	203	
Keldysh Institute of Applied Mathematics	1150	
Mathematica Montisnigri	96	2019-11-13
Matematicheskoe modelirovanie	271	2019-05-21
Keldysh Institute Preprints	884	2021-02-03
Kelompok Peneliti Muda Universitas Negeri Jakarta	32	
Kemala Indonesia Publisher	0	
Kementerian Ketenagakerjaan Republik Indonesia	0	
Kemerovo Institute of Food Science and Technology	0	
Kemerovo State Medical University	269	
Kemerovo State University	1919	

Рис. 1. ИПМ в списке издателей журналов.

На рис. 2 представлен фрагмент алфавитного списка мировых издателей сборников материалов конференций, среди которых ИПМ (Keldysh Institute of Applied Mathematics) представлен материалами двух конференций [4]:

- Futurity designing. Digital reality problems (Проектирование будущего. Проблемы цифровой реальности) – 3 выпуска – 2018, 2019, 2020 гг.
- Scientific Services & Internet (Научный сервис в сети Интернет) – 5 выпусков – 2016 и 2017 гг., а затем объединенные в серию 2018, 2019 и 2020 гг.

На рис. 3 представлен фрагмент алфавитного списка мировых издателей журналов, среди которых КФУ (Kazan Federal University) представлен 11 журналами. Далее, в разделе 6, мы более подробно остановимся на одном из них – журнале «Электронные библиотеки» (Russian Digital Libraries Journal) [5].

В различных инструментах Crossref (при загрузке контента, определении ссылки и т. д.) в зависимости от требований простоты-сложности и оперативности-точности используется тот или иной алгоритм. Но в любом случае автору или сотруднику издательства рекомендуется самостоятельно проверить корректность результата распознавания библиографических записей и точность полученной ссылки.

Publisher	Count 1	Count 2	Icon
Kaunas University of Technology (KTU)	2	2	▶
Kazan Federal University	2	490	▶
Keldysh Institute of Applied Mathematics	10	321	▶

Proceedings	# DOIs		
Proceedings of 18th Scientific Conference "Scientific Services & Internet – 2016"	33		
Proceedings of 19th Scientific Conference "Scientific Services & Internet – 2017"	53		
Series Title: Futurity designing Digital reality problems			
Proceedings of the 1st International Conference "Futurity designing Digital reality problems"	27		
Proceedings of the 2nd International Conference "Futurity designing Digital reality problems"	27		
Proceedings of the 3rd International Conference "Futurity designing Digital reality problems"	24		
Series Title: Scientific Conference "Scientific Services & Internet"			
Proceedings of 20th Scientific Conference "Scientific Services & Internet – 2018"	50		
Proceedings of 21th Scientific Conference "Scientific Services & Internet – 2019"	60		
Proceedings of 22nd Scientific Conference "Scientific Services & Internet – 2020"	45		
Kharkiv National University of Internal Affairs	1	18	▶

Рис. 2. ИПМ в списке издателей сборников материалов конференций.

2. ОТЧЕТ ВКЛАДЧИКА

Возможно, чтобы подчеркнуть тот факт, что издатель, загружая метаданные в Crossref, получает определенную выгоду и некоторые бесплатные сервисы, авторы проекта стали использовать определенные банковские термины: размещение контента называли депонированием, самого издателя – вкладчиком и т. д.

Отчеты вкладчика для каждого издателя используются для проверки основной информации о регистрациях DOI. Отчеты «привязываются» к трем ключевым спискам издателей, поддерживаемым Crossref, – списку издателей журналов, списку издателей сборников конференций и списку издателей монографий (книг) (см. разд. 1). В настоящее время нет отчетов вкладчиков для других типов контента, например, фотографий, картинок, видео и аудио-файлов.

Индексная страница обновляется еженедельно. Отчеты на уровне заголов-

ков обновляются по мере обновления метаданных. Можно получить и проанализировать отчет по всем изданиям ИПМ (журналам, сборникам конференций, монографиям), но мы для примера рассмотрим один из журналов.

Journal	# DOIs	Last Crawl Date
International Trade	1	68
Kazakh-German University	1	102
Kazakhstan Institute for Strategic Studies Under the President of The Republic of Kazakhstan	3	48
Kazan Federal University	12	1656
Eurasian Arabic Studies	1	
Eurasian Arabic Studies	1	
KAZAN LINGUISTIC JOURNAL	82	2021-07-24
Magnetic resonance in solids	57	2021-07-03
Philology and Culture	374	2021-08-09
Uchenye Zapiski Kazanskogo Universiteta Seriya Fiziko-Matematicheskie Nauki	88	
Uchenye Zapiski Kazanskogo Universiteta Seriya Gumanitarnye Nauki	233	2021-08-03
Russian Digital Libraries Journal	194	2019-11-21
Uchenye Zapiski Kazanskogo Universiteta Seriya Estestvennyye Nauki	111	2021-08-07
Tatarica	67	2019-11-19
Izvestiya Vysshikh Uchebnykh Zavedenii Matematika	281	
Education & Self Development	167	2021-07-13
Kazan Law Institute of MIA Russia	1	102
Kazan Medical Journal	0	0
Kazan State Power Engineering University	1	316
Kazan State University of Architecture and Engineering	1	30

Рис. 3. КФУ в списке издателей журналов.

Выбрав в списке журналов (см. рис. 1) издательство Keldysh Institute of Applied Mathematics, а затем, выбрав, например, *Mathematischesкое modelirovanie*, получим подробный отчет, где для каждого DOI указаны префикс владельца, временная метка, дата последнего обновления записи и количество цитирований (по данным Crossref) данной публикации (см. рис. 4).

10.20948/mm-2020-01-01	10.20948	202001101819	2020-01-10	0
10.20948/mm-2020-01-02	10.20948	202001101819	2020-01-10	1
10.20948/mm-2020-01-03	10.20948	202001101819	2020-01-10	2
10.20948/mm-2020-01-04	10.20948	202001101819	2020-01-10	0
10.20948/mm-2020-01-05	10.20948	202001101819	2021-02-21	0
10.20948/mm-2020-01-06	10.20948	202001101819	2020-01-10	2
10.20948/mm-2020-01-07	10.20948	202001101819	2020-01-10	0
10.20948/mm-2020-01-08	10.20948	202001101819	2020-01-10	0
10.20948/mm-2020-02-01	10.20948	202002111347	2020-02-11	0
10.20948/mm-2020-02-02	10.20948	202002111347	2020-02-11	0
10.20948/mm-2020-02-03	10.20948	202002111347	2020-02-11	1
10.20948/mm-2020-02-04	10.20948	202002111347	2020-02-11	0
10.20948/mm-2020-02-05	10.20948	202002111347	2020-02-11	1
10.20948/mm-2020-02-06	10.20948	202002111347	2020-02-11	0
10.20948/mm-2020-02-07	10.20948	202002111347	2020-02-11	2
10.20948/mm-2020-02-08	10.20948	202002111347	2020-02-11	0
10.20948/mm-2020-03-01	10.20948	202002171101	2020-02-17	2
10.20948/mm-2020-03-02	10.20948	202002171101	2020-02-17	0
10.20948/mm-2020-03-03	10.20948	202002171101	2021-02-27	0
10.20948/mm-2020-03-04	10.20948	202002171101	2020-02-17	1
10.20948/mm-2020-03-05	10.20948	202002171101	2020-02-17	0

Рис. 4. Фрагмент Отчета вкладчика для журнала «Математическое моделирование».

3. ОТЧЕТ О КОНФЛИКТЕ

Как известно, DOI — это уникальный идентификатор, поэтому для каждого элемента контента всегда должен быть только один DOI. И издатель получит отчет о конфликтах, если у него есть хотя бы один конфликт с DOI.

Важно исправить эти конфликты как можно скорее, потому что они могут привести к проблемам в будущем. Наличие двух разных DOI для одного и того же контента означает, что исследователь не будет знать, какой из них цитировать, рискуя тем самым исказить количество цитирований. Кроме того, издатель может забыть, что у него есть два DOI на один объект, и обновить только один из них при изменении контента. Это означает, что любой, кто воспользуется другим DOI, который не обновляли, перейдет по неработающей ссылке. Поэтому плохие метаданные следует быстро устранить и тем самым решить проблему.

Отчет о конфликте показывает, где два (или более) DOI были отправлены с одинаковыми метаданными, или указывает на то, что у издателя научной публикации при отправке в Crossref метаданных могли быть повторяющиеся DOI.

Все конфликты DOI, связанные со статьями в журналах, сборниках конференций или монографиях, отмечаются в едином Отчете о конфликте на веб-сайте Crossref (рис. 5). Если у нас есть активные конфликты, мы будем ежемесячно получать напоминание по электронной почте.

Journal	DOI ID	Count	Count	Count
Kasetsart University and Development Institute	10.34044_conflicts.xml	825	1	1
Kastamonu Egitim Dergisi	10.24106_conflicts.xml	4307	1	6
Kaunas University of Technology (KTU)	10.5755_conflicts.xml	45886	3	68
Keldysh Institute of Applied Mathematics	10.20948_conflicts.xml	10274	2	15
Journal		# DOIs		
		0		
Mathematica Montisnigri		15		
Kesmas: Jurnal Kesehatan Masyarakat Nasional	10.21109_conflicts.xml	889	1	1
Kh.Dosmukhamedov Atyrau University	10.47649_conflicts.xml	834	1	1
Kharkiv State Academy of Physical Culture	10.15391_conflicts.xml	2454	1	3
Khayrallah Center for Lebanese Diaspora Studies	10.24847_conflicts.xml	4357	1	6

Рис. 5. Журнал «Математика Черногории» в списке конфликтов DOI.

В отчете о конфликте DOI, представленном на рис. 5, говорится о наличии 15 конфликтных ситуаций, возникших при размещении в Crossref метаданных журнала «Математика Черногории». Первая строчка в этом отчете с неуказанным названием журнала и нулевым количеством конфликтов – это небольшой глюк программистов Crossref, о чем мы им сообщили и с чем они полностью согласились.

А нам нужно разобраться в причине конфликта и постараться его исправить. Кликнув название журнала, получаем полный отчет обо всех 15 конфликтных ситуациях. Пример одной из них:

Created: 2020-03-19 14:23:13.0

ConfID: 5583368

CauseID: 1465359938

OtherID: 1462438643

JT: Mathematica Montisnigri

MD: Jokanović, 46 ,null,5,2019,A breaf survey on Armendariz and central Armendariz rings

DOI: 10.20948/mathmon-2019-46-1(Journal) (5583368-N)

DOI: 10.20948/mathmontis-2019-46-1(Journal)

Как выяснилось, предпринятая редакцией журнала попытка перейти на новые идентификаторы DOI одновременно с использованием старых идентификаторов недопустима с точки зрения идеологии DOI и технологии Crossref.

Как видно из представленного примера, первая статья 46-го тома была загружена дважды с различными идентификаторами (DOI). Хотя для каждого элемента контента всегда должен быть только один DOI, так как наличие двух разных DOI для одного и того же контента может ввести в заблуждение читателей, знакомящихся с материалами данного издания. Да и сами издатели могут запутаться при изменении метаданных и повторной загрузке контента.

Crossref предлагает 3 сценария исправления конфликтов DOI:

Сценарий 1. Если назначили два DOI разным элементам контента, но случайно отправили одни и те же метаданные для них обоих. В этом случае один из DOI имеет неверные метаданные. Если повторно загрузить исправленные метаданные этого DOI, конфликт будет разрешен.

Сценарий 2. Если назначили два DOI одному и тому же элементу контента. В этом случае вы можете разрешить конфликт, назначив один из DOI в качестве основного, а другой — в качестве псевдонима. Псевдоним DOI будет автоматически перенаправляться на основной DOI, поэтому достаточно будет поддерживать только основной.

Сценарий 3. Если два DOI относятся к разным элементам контента, но их метаданные настолько похожи, что был отмечен конфликт. Это происходит, когда в элементы включено очень мало метаданных. Лучше всего зарегистрировать дополнительные метаданные, чтобы устранить конфликт. Или же можно принять конфликт, удалив статус конфликта и установив для него статус «разрешено». Это не повлияет на записи метаданных или DOI, но устранил конфликты из отчета о конфликтах.

Для нас самым простым решением было удалить «неверный» DOI. Но удалять DOI нельзя – это фундаментальный принцип DOI. Поэтому было принято решение идти по второму пути, предложенному Crossref, – «неверный» DOI назвать псевдонимом «правильного» с помощью подсистемы администрирования DOI (рис. 6).

Conflict ID: [5583368](#) generated by submissionID: [1465359938](#)

DOI	Submission	Status	Date	Type	Journal	Version	Year	Volume	Issue	Suppl.	Page	Author	Title	Sequence
10.20948/mathmon-2019-46-1	1462438643	Alias	28-Feb-21 06:25	FULL_TEXT	Mathematica Montisnigni	201912271540	2019	46			5	Jokanovi?	A breaf survey on Amendariz and cent...	
10.20948/mathmontis-2019-46-1	1465359938	Primary	28-Feb-21 06:25	FULL_TEXT	Mathematica Montisnigni	201912271530	2019	46			5	Jokanovi?	A breaf survey on Amendariz and cent...	

Рис. 6. Изменение статуса DOI для разрешения конфликта.

После изменения статуса «неверного» DOI любое обращение к нему вызовет автоматический переход по ссылке, указанной в первичном («правильном») DOI.

Проведя таким образом обновление контента всех «неверных» DOI и изменив их статус на «псевдоним», мы исключили издания ИПМ из списка конфликтов DOI.

4. ОТЧЕТ О ПОЛЯХ ИЛИ ОТСУТСТВУЮЩИХ МЕТАДАНЫХ

Отчет о полях или отсутствующих метаданных содержит подробную информацию о полноте метаданных. Он так же, как и «Отчет вкладчика», «привязывается» к одному из трех ключевых списков издателей, поддерживаемых Crossref, – списку издателей журналов, списку издателей сборников конференций и списку издателей монографий (книг) (см. разд. 1).

К данному отчету можно получить доступ, выбрав значок (зеленая стрелка вправо) рядом с именем издателя научных материалов в одном из вышеуказанных списков.

Следует отметить, что наборы метаданных, передаваемых для статей журналов, научных публикаций в сборниках конференций и монографий, несколько различаются. Для любой научной публикации передаются ее название на русском и английском языках, список авторов (для каждого автора имя (First name) и фамилия (Surname) на русском и английском языках, а также ORCID), год издания, количество или диапазон страниц. Для журналов и сборников добавляются номер тома и/или номер выпуска. Кроме того, последние 2 года мы стали за-

гружать в Crossref аннотации научных публикаций и списки литературы, что очень важно для отслеживания взаимного цитирования [3].

Рассмотрим информацию по нашим журналам, представленную в отчетах о полях или отсутствующих метаданных (рис. 7).

REPORT DATE:

Publication Title	Ignore Fields	V	I	P	A	S	V/I	V/I/P	V/I/P/A	P/A	T	N	F	missing iParadigms URLs	Total DOIs*
Keldysh Institute Preprints	none CHANGE	878	0	0	0	275	0	0	0	0	0	0	0	na	878
Matematicheskoe modelirovanie	none CHANGE	25	0	0	0	29	0	0	0	0	0	0	25	na	162
Mathematica Montisnigri	none CHANGE	0	89	0	2	21	0	0	0	0	0	2	1	na	89

Рис. 7. Отчет о полях или отсутствующих метаданных в журналах ИПМ.

Заголовки столбцов таблицы на рис. 7 (красным отмечены поля, с точки зрения Crossref неполные или некорректно заполненные):

- Participation Title – название издания;
- Ignore Fields – указание, отсутствие каких полей следует игнорировать;
- V=volume – номер тома;
- I=issue – номер выпуска;
- P=page – количество или диапазон страниц;
- A=author – автор;
- S=single-author – единственный автор;
- T=articletitle – название статьи;
- N=no-first-name – не задано имя автора;
- F=first name initial only – в качестве имени заданы только инициалы автора;
- U=missing iParadigms Url– Crossref сотрудничает (с 2008 г.) с iParadigms LLC, предлагая своим членам — ведущим научным и профессиональным издателям — возможность проверки оригинальности работ с помощью служб CrossCheck и iThenticate. База данных CrossCheck включает полнотекстовые журналы ведущих академических издателей и достаточно быстро растет по мере того, как издатели-участники Crossref подписываются на эту услугу.

Как мы видим, с помощью Отчета о полях или отсутствующих метаданных Crossref подчеркивает – несмотря на то, что некоторые библиографические ме-

таданные являются необязательными для целей регистрации контента, настоятельно рекомендуется всем издателям регистрировать максимально полные метаданные для каждого зарегистрированного элемента. И даже выделяет красным цветом в отчете поля, для которых, с точки зрения Crossref, издатель грубо не выполняет указанные рекомендации.

Анализируя информацию, представленную на рис. 7, можно отметить, что для всех выпусков Препринтов ИПМ не заданы номера тома, для всех томов журнала «Математика Черногории» не заданы номера выпусков. Для журнала «Математическое моделирование» замечания касаются не всех выпусков, а только нескольких томов начального периода присвоения DOI научным публикациям, когда нашему журналу присваивали DOI какие-то внешние организации, даже некорректно зарегистрированные в Crossref.

Следует отметить, что всем статьям журнала «Математика Черногории» DOI присваивали мы. И, соответственно, можем вносить необходимые исправления и/или дополнения в наборы метаданных, загруженные в Crossref. С журналом «Математическое моделирование» ситуация более сложная – до 2020 года в результате конкурсов издание переходило из рук в руки и получало совершенно различные DOI. При этом отгружались метаданные научных статей журнала, абсолютно различные по полноте и корректности. И теперь ИПМ, как издатель журнала, не имеет ни прав, ни возможностей что-либо исправить или дополнить в информации о выпусках 2016–2019 гг.

Анализируя далее замечания Crossref по регистрации метаданных публикаций, мы видим отметки в столбце задания единственного автора (S=single-author) и некоторые другие.

Похожие замечания имеются по метаданным сборников конференций и монографий. Мы с этими (и другими) замечаниями Crossref внимательно разбираемся и по возможности стараемся устранить, внося необходимую правку в описание метаданных.

Однако издателю конкретного журнала можно отказаться от подобной навязчивости Crossref с помощью переключателя Change (второй столбец, рис. 7 – IgnoreFields), указав, отсутствие каких полей следует игнорировать (рис. 8).

data.crossref.org/ignoreFields/?jciteid=376604

← Back to the main Crossref website

Crossref

Field Report - Ignore Field(s) Request

Select the fields you would like ignored in future reports for:
Journal of NBC Protection Corps

Volume Issue
 Page Author
 Article Title First Name Initial Only
 Single Author Only No First Name

Submit Cancel

Рис. 8. Указание, отсутствие каких полей следует игнорировать.

5. ОТЧЕТ СКАНЕРА DOI

Отчет сканера DOI выполняется только для журналов и, соответственно, «привязывается» к списку издателей журналов.

Сканер DOI делает выборку статей для каждого журнала конкретного издателя, чтобы убедиться, что заданные DOI переводятся на соответствующую страницу. Для каждого просматриваемого журнала выбираются DOI, количество которых равно примерно 5% от общего числа DOI для журнала (максимум до 50 DOI). На рис. 9 перечислены все журналы, издаваемые ИПМ, и для каждого указаны общее количество DOI и дата последнего сканирования.

Journal	# DOIs	Last Crawl Date
Mathematica Montisnigri	96	2019-11-13
Matematiccheskoe modelirovanie	271	2019-05-21
Keldysh Institute Preprints	884	2021-02-03

Рис. 9. Дата последнего сканирования журналов ИПМ.

Получить доступ к деталям работы поискового робота для данного журнала можно, выбрав дату в столбце «Дата последнего сканирования».

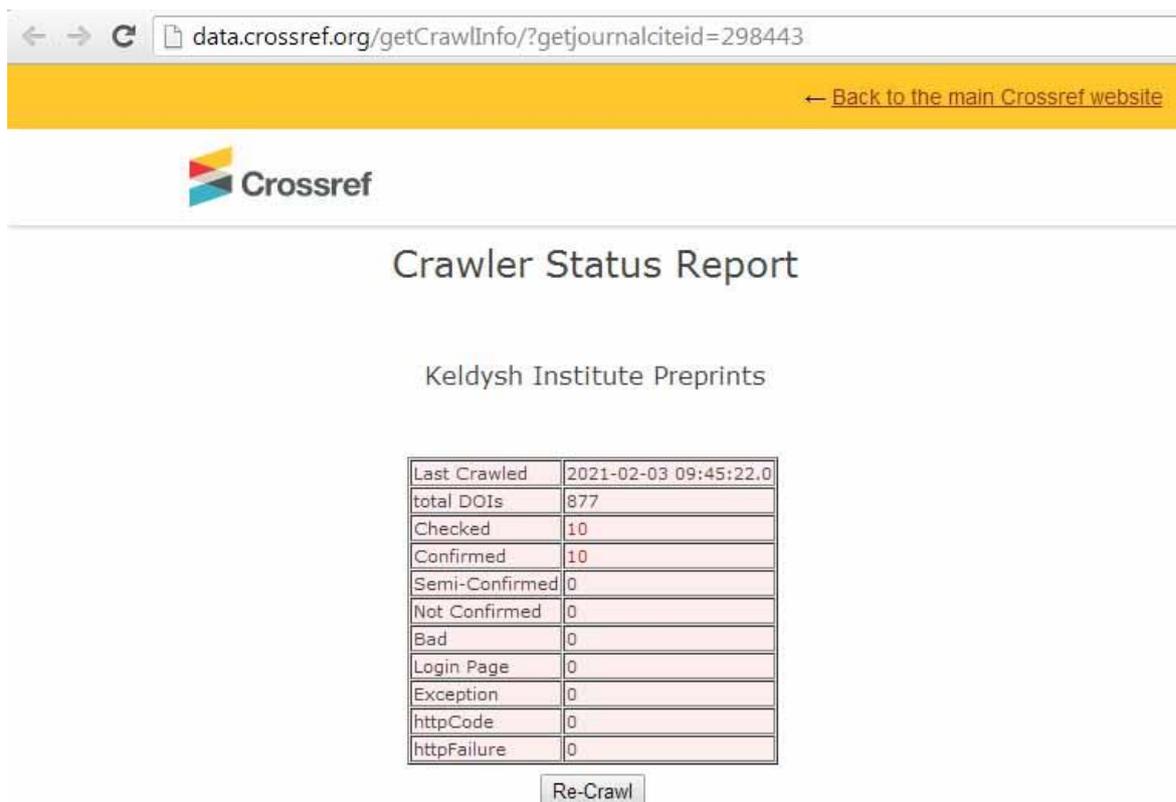


Рис. 10. Отчет о работе сканера DOI для препринтов ИПМ.

Никаких ошибок, как следует из отчета сканера на рис. 10, у Препринтов ИПМ нет. При этом просканировано (и подтверждена корректность) 10 DOI. Эти поля «кликабельны», поэтому можно посмотреть более подробно результаты процедуры сканирования (рис. 11, 12).

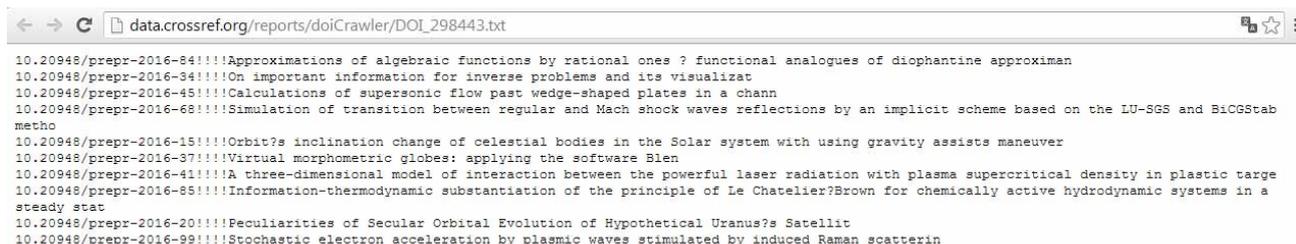
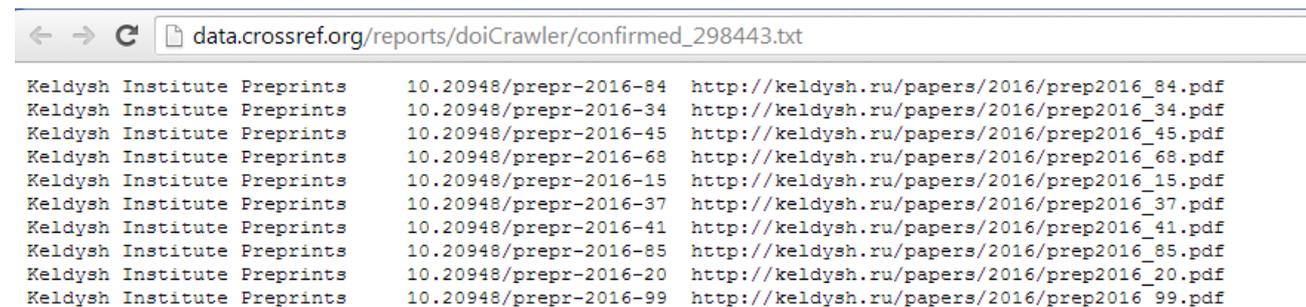


Рис. 11. Список выбранных для сканирования DOI.



Source	DOI	URL
Keldysh Institute Preprints	10.20948/prepr-2016-84	http://keldysh.ru/papers/2016/prep2016_84.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-34	http://keldysh.ru/papers/2016/prep2016_34.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-45	http://keldysh.ru/papers/2016/prep2016_45.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-68	http://keldysh.ru/papers/2016/prep2016_68.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-15	http://keldysh.ru/papers/2016/prep2016_15.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-37	http://keldysh.ru/papers/2016/prep2016_37.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-41	http://keldysh.ru/papers/2016/prep2016_41.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-85	http://keldysh.ru/papers/2016/prep2016_85.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-20	http://keldysh.ru/papers/2016/prep2016_20.pdf
Keldysh Institute Preprints	10.20948/prepr-2016-99	http://keldysh.ru/papers/2016/prep2016_99.pdf

Рис. 12. Проверенные и «подтвержденные» DOI.

Эта функция (сканирование DOI), на наш взгляд, достаточно интересная и полезная, особенно для тех изданий, у которых изменялись сервер базы данных или администрация, так как такого рода изменения могли внести серьезные ошибки в процедуру депонирования контента и сделать недоступными материалы изданий.

Однако реализация этой функции программистами Crossref представляется немного странной и вызывает несколько вопросов:

- Почему, например, для Препринтов ИПМ, насчитывающих 877 выпусков с начала присвоения DOI, проверяется только 10 элементов, хотя в описании алгоритма говорится о 5% от общего числа?
- Почему при повторном сканировании (Re-Crawl) используется та же выборка, что и несколько лет назад, при предыдущей проверке? Ведь издателю интересно отслеживать корректность не только старых материалов, но и всех публикаций исследуемого журнала.

Пока что мы не получили ответа на эти вопросы от Службы поддержки Crossref, но продолжаем активное взаимодействие с ней.

Еще одно важное дополнение связано с тем, что Crossref – это один из регистраторов DOI. А так как регистраторы DOI не обмениваются метаданными, то издатели, связанные с другими регистрационными сервисами, не смогут воспользоваться инструментами Crossref.

6. ОТЧЕТ ОБ УЧАСТИИ

Авторы проекта и разработчики Crossref призывают издателей не просто размещать метаданные научных публикаций, а делать их максимально полными. Кроме того, Crossref призывает научное и издательское сообщество активно

пользоваться предлагаемыми сервисами, разработанными для анализа полноты и корректности загружаемой информации, тем самым как бы участвуя в развитии и расширении набора этих услуг.

Для каждого издателя, сотрудничающего с Crossref, существует отдельный Отчет об участии (Participation report), который показывает, какой процент их депонированных данных зарегистрирован для каждого из десяти ключевых элементов метаданных. Отчеты об участии наглядно показывают, где есть проблемы и что можно улучшить в плане полноты метаданных.

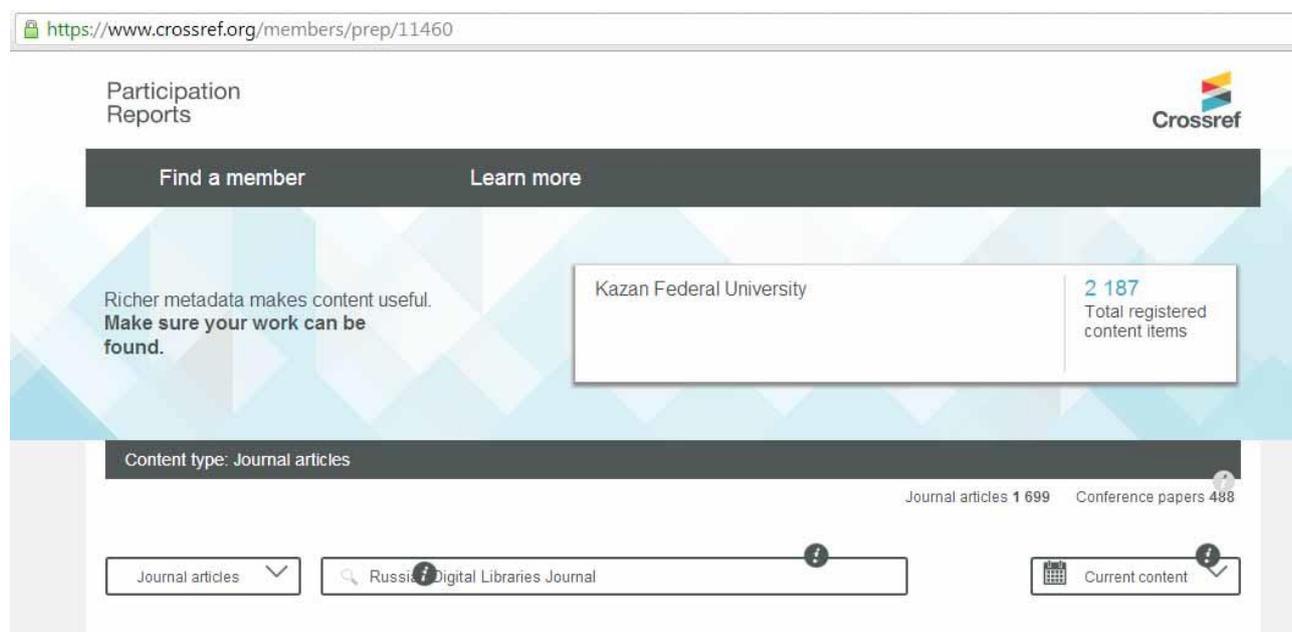


Рис. 13. Отчет об участии для научных публикаций КФУ.

На рис. 13 представлен заголовок Отчета об участии издателя Kazan Federal University – общее число элементов контента 2187, в том числе журнальных статей 1699, статей в сборниках материалов конференций 488. В нижней части заголовка (рис. 13) есть 2 меню – выбор типа научной публикации (слева) и выбор анализируемого периода (все время депонирования, текущий период – последнее 2 года, «старые» материалы – данные, загруженные более 2 лет назад). Центральное поле заголовка Отчета об участии позволяет ввести название журнала, сборника или даже название публикации и проанализировать полноту соответственно загруженных метаданных.

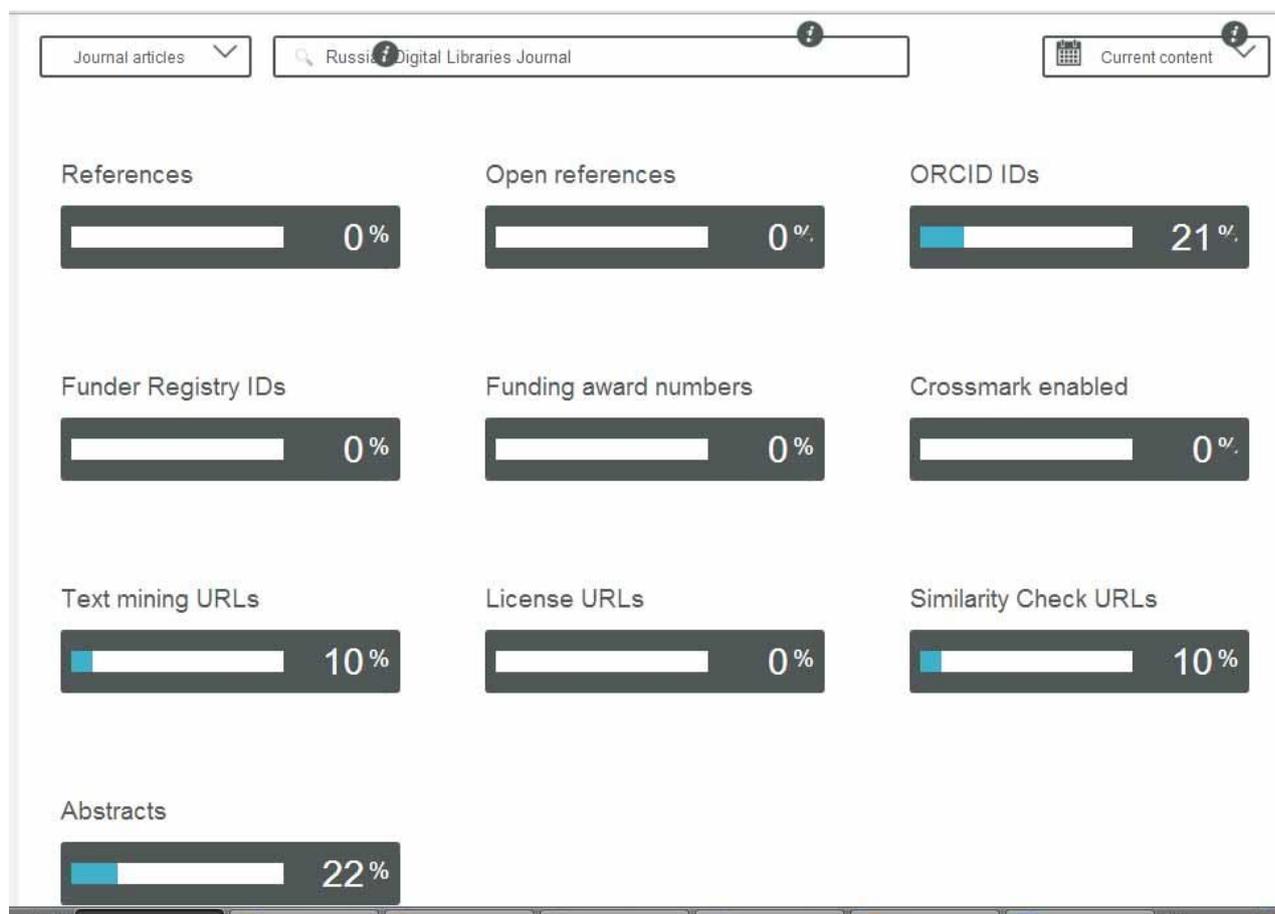


Рис. 14. Процентные показатели Отчета об участии журнала «Электронные библиотеки» (“Russian Digital Libraries Journal”).

Основная часть Отчета об участии для журнала «Электронные библиотеки» показана на рис. 14 (задан текущий период (current time) – последние 2 года):

- для 0 % публикаций загружен список литературы (References);
- 0 % ссылок открыты (Open References) – доступны всем пользователям всех сервисов Crossref (нет доступных ссылок, так как нет списков литературы);
- для 21 % авторов указан ORCID;
- для 0 % публикаций указаны имя и идентификатор (Funder Registry IDs) спонсора – хотя бы одной из организаций, финансировавших исследование;
- для 0 % публикаций указан номер гранта финансирования (Funding award numbers);
- доля контента (в данном случае 0 %), использующего службу Crossmark (Crossmark-enabled), которая дает читателям быстрый и легкий доступ

к текущему статусу элемента контента (в рамках политики издателя в отношении исправлений, опровержений, отзыва и других обновлений);

- процент зарегистрированного контента (в данном случае 10 %), содержащего URL-адреса для интеллектуального анализа текста и данных (Text-mining URLs) научной публикации – автоматического анализа и извлечения информации из большого количества документов. В настоящий момент большинство научных организаций мира (и КФУ, как нам кажется, в том числе) не заинтересованы в задании специального набора инструкций, с помощью которых кто-то зачем-то будет исследовать их научные материалы;
- процент метаданных публикаций (в данном случае 0 %), содержащих URL-адреса, указывающие на лицензию (License URLs), определяющую условия, на которых читатели могут получить доступ к контенту;
- процент метаданных публикаций (в данном случае 10 %), которые включает URL-адреса для проверки схожести (Similarity Check URLs), для изданий, сотрудничающих с CrossCheck и iThenticate;
- 22 % метаданных публикаций включают аннотации (Abstracts), что дает более глубокое понимание содержания работы.

На наш взгляд, не стоит гнаться за 100 % показателями, но при этом должно быть понятно, что более полное и аккуратное заполнение метаданных публикации в той или иной мере влияет [6] на рейтинги изданий, авторов и организаций. А указание грантов и фондов поддержки научной деятельности положительно влияет на взаимоотношения с этими фондами.

СПИСОК ЛИТЕРАТУРЫ

1. Ассоциация Crossref. URL: <https://www.crossref.org/about/>
2. International DOI Foundation (IDF). URL: <https://www.doi.org/>
3. *Ермаков А.В.* Библиографическая ссылка как инструмент автора и читателя // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции (21–25 сентября 2020 г., онлайн). М.: ИПМ им. М.В. Келдыша, 2020. С. 268–275. <https://doi.org/10.20948/abrau-2020-55>
URL: <https://keldysh.ru/abrau/2020/theses/55.pdf>
4. *Слепенков М.И.* Материалы конференций в онлайн-библиотеке ИПМ им. М.В. Келдыша // Препринты ИПМ им. М.В. Келдыша. 2020. № 18. 16 с.

<http://doi.org/10.20948/prepr-2020-18>

URL: <http://library.keldysh.ru/preprint.asp?id=2020-18>

5. Журнал «Электронные библиотеки». URL: <https://elbib.ru>

6. *Полилова Т.А.* Инфраструктура научных публикаций // Препринты ИПМ им. М.В. Келдыша. 2009. № 15. 30 с.

URL: <http://library.keldysh.ru/preprint.asp?id=2009-15>

IMPROVING THE QUALITY OF METADATA SCIENTIFIC PUBLICATIONS WITH CROSSREF REPORTS

Alexey V. Ermakov^[0000-0002-6054-0813]

Ermakov@Keldysh.ru

Keldysh Institute of Applied Mathematics, Miusskaya sq., 4, Moscow, 125047, Russia

Abstract

Issues related to improving the quality of metadata of scientific publications placed in the Crossref bibliographic database are considered. All information contained in metadata obtained from publishers of scientific publications, Crossref analyzes and displays in various reports. The reports give publishers an idea of the completeness and correctness of the bibliographic data provided. The quality of metadata directly or indirectly affects the number of views and links to a publication, respectively, on the ratings of scientific publications, authors and organizations.

Keywords: *metadata of publications, Crossref reports, citations, ratings of scientific publications.*

REFERENCES

1. Crossref Association. URL: <https://www.crossref.org/about/>
2. International DOI Foundation (IDF). URL: <https://www.doi.org/>
3. *Ermakov A.V.* Bibliographic reference as a tool for the author and reader // Conference material: “Scientific service & Internet: proceedings of the 22nd All-Russian Scientific Conference (September 21-25, 2020, online)”. M.: KIAM
URL: <https://keldysh.ru/abrau/2020/theses/55.pdf>

4. *Slepenkov M.I.* Conference proceedings in the on-line library of Keldysh Institute // KIAM Preprint № 18, Moscow, 2020. 16 p.

<http://doi.org/10.20948/prepr-2020-18>

URL: <http://library.keldysh.ru/preprint.asp?id=2020-18>

5. Russian Digital Libraries Journal. URL: <https://elbib.ru>

6. *Polilova T.A.* Infrastructure of scientific publications // KIAM Preprint № 15, Moscow, 2009. 30 p.

URL: <http://library.keldysh.ru/preprint.asp?id=2009-15>

СВЕДЕНИЯ ОБ АВТОРЕ



Ермаков Алексей Викторович – старший научный сотрудник Института прикладной математики им. М.В. Келдыша Российской академии наук

Alexey V. Ermakov – Senior Researcher, Keldysh Institute of Applied Mathematics.

email: Ermakov@Keldysh.ru,

ORCID: 0000-0002-6054-0813

Материал поступил в редакцию 24 октября 2021 года