

<https://doi.org/10.24108/2658-3143-2020-3-2-3-155-165>



ОРИГИНАЛЬНЫЕ СТАТЬИ / ORIGINAL ARTICLES

Многоликий DOI CrossRef: все ли функции мы используем?

Наталья Н. Литвинова^{1,2}

¹Некоммерческое партнерство «Национальный электронно-информационный консорциум» (НЭИКОН)
ул. Летниковская, 4, стр. 5, офис 2.4, г. Москва, 115114, Российская Федерация

²Российская государственная библиотека
ул. Воздвиженка, 3/5, г. Москва, 190019, Российская Федерация

Аннотация

В статье рассматриваются новые функции цифровых идентификаторов объектов — Digital Object Identifiers (DOI), регистрируемых международным агентством CrossRef. Эти функции развиваются на основе более чем стомиллионной базы идентификаторов с развитым набором метаданных. Показано, какую роль играет база DOI CrossRef в отслеживании событий упоминания публикаций помимо цитирования в списках литературы и в вычислении показателей альтметрик. Проанализировано значение DOI CrossRef в определении статуса свободного доступа к публикациям с помощью сервиса Unpaywall. Выделяется несколько типовых проблем российских публикаций, которые препятствуют корректному определению статуса свободного доступа: отсутствие DOI, пробелы в представлении метаданных в CrossRef и на веб-страницах публикаций, ошибки при регистрации DOI.

Ключевые слова: DOI, CrossRef, цифровой идентификатор объекта, списки литературы, ссылки, упоминания публикаций в Интернете, альтметрики, свободный доступ, Unpaywall

Благодарности: автор благодарит Ричарда Орра (Our Research, Unpaywall) за предоставленные разъяснения о работе Unpaywall и Д.В. Косякова (ГПНТБ СО РАН) за содержательные комментарии и дополнения к тексту рукописи.

Для цитирования: Литвинова Н. Н. Многоликий DOI CrossRef: все ли функции мы используем? *Наука и научная информация*. 2020;3(2-3):155-165. <https://doi.org/10.24108/2658-3143-2020-3-2-3-155-165>

Статья поступила: 13.08.2020

Статья принята в печать: 24.09.2020

Статья опубликована: 15.11.2020

Multifaceted CrossRef DOI: Do We Use All Functions?

Natalia N. Litvinova^{1,2}

¹Non-profit Partnership “National Electronic Information Consortium” (NEICON)
Letnikovskaya str., 4, bldg 5, off. 2.4, Moscow, 115114, Russia

²Russian State Library
Vozdvizhenka str., 3/5, Moscow, 190019, Russia

Abstract

The article deals with some new functions of DOI CrossRef system. Based on current database of more than 100 million of records containing DOIs and advanced sets of metadata new important functions of DOIs emerged. DOI is now widely used for discovery and accounting of publications' mentions on the Internet by altmetric services and CrossRef DOI Event Data. DOI CrossRef is also served as a basis for Unpaywall service which defines open access status of publications. Analyzed are typical problems of Russian publications preventing correct discovery of open access evidence by Unpaywall: missing DOIs, incomplete metadata registered with CrossRef and on publications web pages on the Internet, mistakes in DOIs.

Keywords: DOI, CrossRef, digital object identifier, reference lists, links, mentions, altmetrics, open access, Unpaywall

Acknowledgements: the author thanks Richard Orr (Our Research, Unpaywall) for his support in explaining details of Unpaywall analysis and Denis Kosiakov (SPSTL SB RAS) for his comments and suggestions on the manuscript.

For citation: Litvinova N. N. Multifaceted CrossRef DOI: Do We Use All Functions? *Scholarly Research and Information*. 2020;3(2-3):155-165. <https://doi.org/10.24108/2658-3143-2020-3-2-3-155-165>

Received: 13.08.2020

Revised: 24.09.2020

Published: 15.11.2020

1. Введение

Цифровые идентификаторы объектов — Digital Object Identifiers (DOI) играют все более важную роль в современной научной коммуникации. Цифровые идентификаторы публикаций, зарегистрированные в CrossRef с корректными метаданными, позволяют решать важную задачу установления связей между автором и его публикациями, между библиографическими ссылками на публикации и адресами их местонахождения в электронной среде. База международного регистрационного агентства CrossRef, регистрирующего DOI с 2000 года, в настоящее время насчитывает более 100 миллионов DOI¹. На ее основе развиваются новые функции DOI CrossRef. Основная задача настоящей статьи — привлечь внимание российского научно-информационного сообщества к тем функциям, которые, по мнению автора, не очень широко известны. Как следствие, не осознается ряд проблем, возникающих из-за недостаточной распространенности DOI CrossRef в российских публикациях, бедности метаданных, регистрируемых в CrossRef, и отсутствия формализованного оформления DOI CrossRef на страницах российских электронных изданий. Решение этих проблем очень важно для углубленной интеграции российских публикаций в международную инфраструктуру научных коммуникаций и повышения комфортности работы пользователей с базами научных публикаций. В рамках решения поставленной задачи автор постарался описать типовые проблемы и проиллюстрировать их яркими примерами «потерь» российских публикаций в альтметрических

показателях и в определении статуса свободного доступа к ним с помощью сервиса Unpaywall. Следующим этапом может стать количественный анализ масштабов этих потерь.

2. DOI CrossRef как сервис обеспечения перехода по ссылкам в электронных публикациях

Цифровые идентификаторы объектов DOI появились в результате совместных усилий зарубежных издателей по решению проблемы связывания публикаций различных контент-провайдеров. В 1998 году был создан фонд IDF — International DOI Foundation, целью которого стало создание и продвижение системы DOI как инфраструктуры управления онлайн-контентом [1]. В 2000 году при участии IDF и нескольких ведущих западных издательств был учрежден CrossRef — сервис обеспечения ссылок (linking service) ассоциации PILA — Publisher International Linking Association. Как явствует из названия ассоциации, функциональным назначением DOI CrossRef было обеспечение надежного перехода по ссылкам от кратких библиографических записей, прежде всего приведенных в списках литературы (references), к детальным описаниям документов на платформах контент-провайдеров и к их полным текстам при наличии прав доступа. Эту функцию DOI успешно выполняет уже более двадцати лет. При регистрации DOI в CrossRef регистрирующая организация передает туда метаданные документа, включающие обязательные и опциональные элементы метаданных; в последние входят и списки литературы (которые в профес-

¹ В настоящее время наряду с CrossRef функционируют и другие регистрационные агентства, присваивающие DOI на национальном уровне (например, китайский CNKI) или присваивающие DOI определенным типам научной продукции (например, DataCite).

сиональной литературе не всегда относят к метаданным документов). По метаданным записей этих списков CrossRef ищет DOI и возвращает организации не только DOI переданного документа, но и DOI элементов списка литературы, для которых идентификаторы были найдены. С 2017 года CrossRef рекомендует записывать DOI как URL, разрешающийся в страницу с описанием документа. Эта рекомендация опубликована на сайте CrossRef в DOI Display Guidelines (<https://doi.org/10.13003/5jchdy>).

Синтаксис URL имеет вид:

<https://doi.org/10.xxxx/xxxxx>,

где 10. — стандартный элемент префикса DOI, xxxx до слеша — идентификатор организации (как правило, издательства), зарегистрированной в CrossRef, xxxxx после слеша — суффикс: идентификатор документа, присвоенный ему организацией по своим внутренним правилам.

Благодаря этим устоявшимся процедурам большинство публикаций современных западных издательств содержат в списках литературы надежные ссылки, позволяющие перейти к цитируемому документу за один-два клика, что существенно облегчает работу исследователей по изучению литературы по интересующей их теме. Сравнивая эту ситуацию с поиском документов, упомянутых в списках литературы в эпоху «до Интернета», тогдашний исполнительный директор CrossRef Ed Penz пишет, что прежние поиски часто оканчивались «библиографической фрустрацией» [2, с. 207].

На момент начала использования DOI существовали и другие методы обеспечения ссылок, в частности метод ссылок по OpenURL. Суть его — в передаче метаданных ресурса-источника (например, элемента списка литературы) на вход специального сервиса link resolver, который возвращает адреса документов, обладающих заданным набором метаданных. Если ссылка по DOI всегда ведет на документ на платформе издателя² через центральный сервис, то ссылка по OpenURL «опрашивает» множество платформ и предоставляет возможность выбора версии документа на разных платформах: издателя, агрегаторов, репозиториев. Кроме того, link resolver дает возможность библиотекам настроить его работу таким образом, чтобы пользователю предъявлялись только ссылки на платформы, к которым у библиотеки есть доступ. Являясь в определенном смысле конкурента-

ми, ссылки по DOI и OpenURL сосуществуют, дополняя и страхуя друг друга, причем DOI включается в набор элементов метаданных, которыми оперирует link resolver.

В начале двухтысячных ссылки по DOI получали далеко не все библиографические записи из списка литературы, поскольку публикации до 2000 года просто не имели DOI. В работе [4] приводится пример статьи 2001 года из журнала Journal of Physics A — “Calculation of Lyapunov exponent using an equivalent stochastic system”, — в которой из 13 элементов списка литературы только один имел ссылку по DOI, в то время как ссылки по OpenUR присутствовали во всех элементах списка. Сейчас в этой статье уже 10 элементов списка имеют ссылку по DOI, даже на публикации 1970-х годов. Издатель, Institute of Physics, заказал ретроспективное присвоение DOI для всех публикаций на своей платформе. Так же поступили большинство крупных западных издателей. Это убедительно свидетельствует о важности DOI для современной научной коммуникации: издатели готовы платить за DOI, понимая, что удобство ссылочного аппарата увеличивает привлекательность их продуктов для пользователей, а значит, и их стоимость.

3. Развитие структуры метаданных, регистрируемых в CrossRef

В начале своей деятельности CrossRef ограничивался минимальным набором метаданных для регистрации DOI. В число обязательных элементов не включалось даже заглавие публикации [5, 6]. С годами полнота и точность метаданных приобретала все большее значение. Конференция CrossRef 2018 года “How good is your metadata” была посвящена качеству метаданных, поставляемых в базу CrossRef. Сейчас в универсальных метаданных Crossref (применимых ко всем типам контента) выделяются базовые (заглавие, автор, даты публикации, название источника, том, выпуск, страницы, электронный адрес) и расширенные, к числу которых относятся рефераты/аннотации, списки использованной литературы, информация о лицензиях, фондах, аффилиации и ORCID авторов, связях с другими публикациями и другие. Причем количество записей, содержащих расширенный набор метаданных, устойчиво растет. Специальный сервис Participation Reports с визуализацией (<https://www.crossref.org/members/prep/>) позволяет получить картину расширенного состава метаданных в ресурсах

² До недавнего времени ссылка по DOI разрешалась строго в один документ на платформе издателя. Однако сейчас в некоторых случаях стало допускаться множественное разрешение DOI [3].

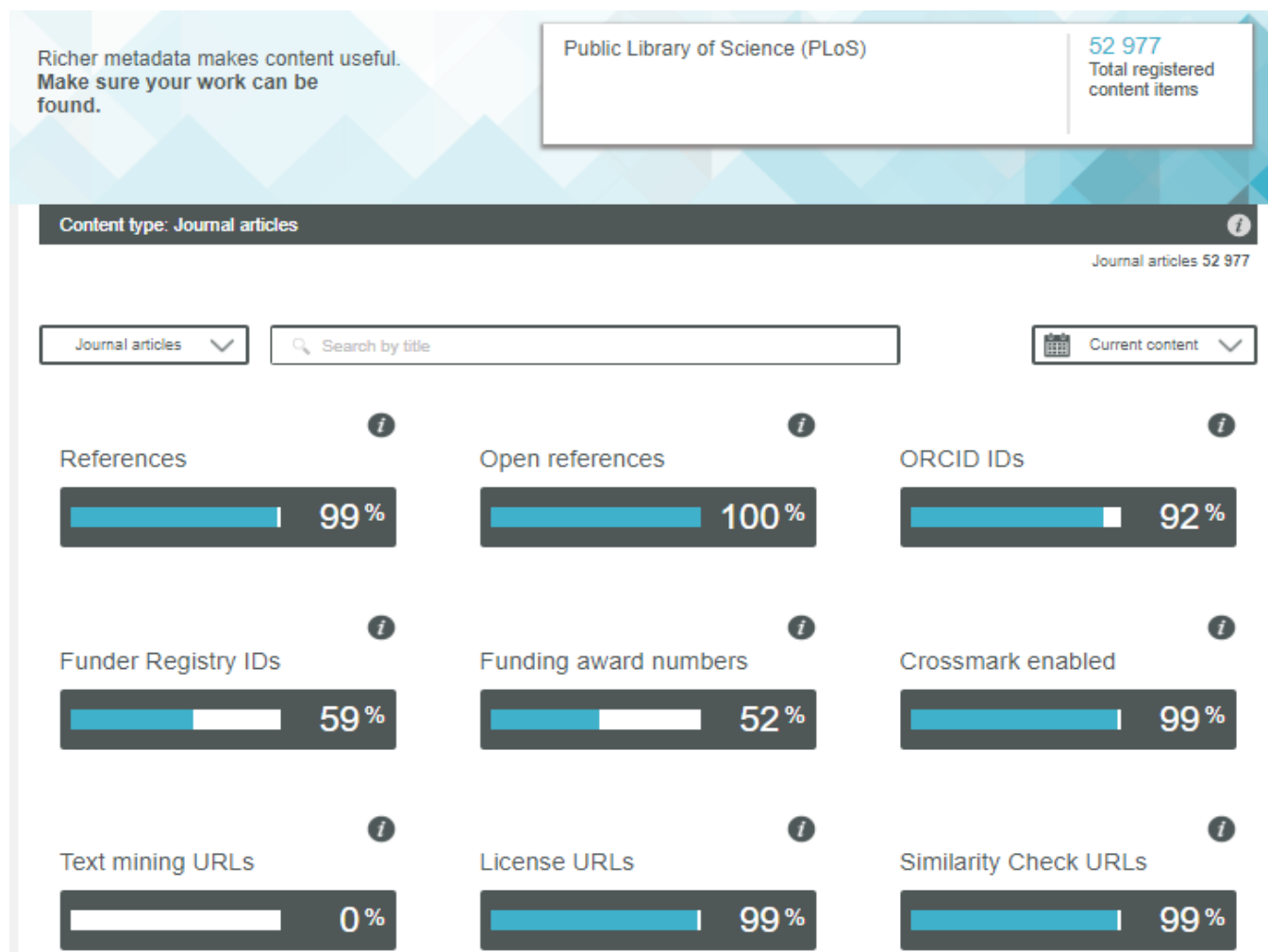


Рис. 1. Participation Report для издательства Public Library of Science

Fig. 1. Participation Report for Public Library of Science

организации-члена CrossRef, как показано на рисунке 1. Кроме того, CrossRef обогащает метаданные, переданные организацией при регистрации DOI, дополняя их новыми выявленными связями (например, между оригиналом и переводом, статьей и набором данных), информацией о фондах, классификационными кодами журналов Scopus и другими. Осознавая важность расширенного состава метаданных для поиска публикаций и анализа публикационных процессов, издатели не только включают их в текущие регистрации, но и активно добавляют в ранее сформированные метаданные. Как показано в [6], по состоянию на 2019 год метаданные, зарегистрированные в период с 2013 по 2016 год, подверглись корректировке в 80% случаев.

4. Новые функции DOI CrossRef

В настоящее время CrossRef насчитывает более 100 миллионов зарегистрированных DOI. Как было показано выше, DOI присваиваются не только текущим публикациям, но и вышедшим десятки и сотни лет назад. Каждому DOI сопоставлен набор метаданных, включающий базовые и все чаще — расширенные метаданные. Это закономерно приводит к развитию новых опций DOI помимо первоначально планировавшихся функций однозначной идентификации объектов и связывания элементов списков использованной литературы с элементами на издательских платформах. Можно выделить по крайней мере три новые функции DOI CrossRef:

- база для анализа публикационных потоков и публикационной активности;

- основа работы сервисов отслеживания упоминаний публикаций в широком контексте сетевых публикаций: собственный сервис CrossRef Event Data, сервис Altmetric и подобные;
- основа работы сервиса определения статуса свободного доступа к публикациям Unpaywall.

4.1. Анализ публикационных потоков и публикационной активности

База метаданных CrossRef объемом более 100 миллионов записей представляет собой ценнейший источник данных для библиометрического и наукометрического анализа, включая анализ цитирования, но не ограничиваясь им. В [6] приводятся многочисленные примеры работ, в которых выполнен многоаспектный анализ публикаций на базе метаданных CrossRef. Данные из базы можно забирать с помощью API (<https://www.crossref.org/education/retrieve-metadata/rest-api/>) без ограничений по контенту: метаданные не охраняются лицензией, они распространяются свободно согласно законодательству об авторском праве.

4.2. CrossRef Event Data и сервисы альтметрик

Альтернативные метрики, или альтметрики, относятся к упоминаниям публикаций в контекстах, отличных от цитирования, которое в рамках этого противопоставления рассматривается как объект традиционных метрик. Упоминания (mentions) отслеживаются в социальных сетях общего назначения, таких как Twitter и reddit, научных социальных сетях (Mendeley), новостных агрегаторах, блогах и других. Эти упоминания становятся все более важными показателями влияния публикаций, выигрывая у цитирования по оперативности и дополняя показатели цитирования характеристиками новых аспектов бытования публикаций в Сети [7, 8].

Первым сервисом, отслеживающим и подсчитывающим упоминания, стал возникший в 2011 году Altmetric.com (<https://www.altmetric.com/>). Он ориентирован на издательский мир; этот сервис интегрирован во множество издательских платформ, для которых поставляются данные об упоминаниях публикаций на коммерческой основе. Помимо этого, Altmetric.com предлагает бесплатный букмарклет (браузерная закладка, выполняющая небольшой JavaScript-код, указанный в ее URL-адресе) Altmetric it!, который может вывести данные альтметрик на веб-странице публикации. Основным условием работы Altmetric it! является наличие DOI. Если на странице публикации нет DOI или сервис не может его опознать как идентификатор данной публикации, появляется сообщение: “Sorry!

We couldn't find a DOI on this page” — «Извините, мы не можем найти DOI на этой странице». Существует также список зарегистрированных в сервисе журналов, для которых можно получить информацию об упоминаниях их публикаций даже при отсутствии DOI,

CrossRef Event Data, появившийся в 2017 году, также отслеживает упоминания в Сети публикаций, имеющих DOI. Но в отличие от Altmetric.com и других сервисов альтметрик (например, PlumX), Event Data не подсчитывает значения метрик, а генерирует по запросам пользователей через бесплатный API информацию обо всех событиях упоминания для определенного DOI или совокупности DOI, имеющих общий префикс (разумеется, в источниках, отслеживаемых этим сервисом). Их девиз: “We provide the unprocessed data — you decide how to use it” — «Мы предоставляем необработанные данные — вы решаете, как их использовать» (<https://www.crossref.org/services/event-data/>). CrossRef Event Data — развивающийся сервис; после периода бета-тестирования разработчики продолжают активно его совершенствовать.

Данные поставляются в файле json, из которого можно извлечь информацию об источниках упоминания, времени упоминания, адресе (URL) страницы, упоминающей публикацию, и другие данные. Например, запрос об упоминаниях препринта “The Future of OA: A large-scale analysis projecting Open Access publication and readership” [9] с начала 2020 года оформляется так:

```
https://api.eventdata.crossref.org/v1/events?
mailto=your e-mail&obj-id=10.1101/795310&from-
occurred-date=2020-01-01,
```

где your e-mail — адрес электронной почты автора запроса.

По этому запросу получаем данные в json, из которых извлекаем информацию о 26 упоминаниях этой публикации с начала 2020 года, среди которых 4 — в научных блогах, 2 — в Твиттере, два — на Zenodo и 18 — в ссылках из статей франкоязычной Википедии.

CrossRef Event Data, несомненно, представляет интерес для отслеживания упоминаний как отдельных публикаций, так и публикаций определенных журналов и издательств, имеющих индивидуальные префиксы DOI. Эта информация поможет выявить каналы продвижения публикаций и группы заинтересованных пользователей. Но для полноценного срабатывания Event Data нужно, чтобы публикации имели DOI, а в социальных сетях и других сетевых сообществах утвердилась

практика включения DOI в упоминания публикаций. Можно констатировать, что такая ситуация уже имеет место в странах «условного Запада», но еще далека от этого состояния в России: ссылки на публикации в социальных сетях редко включают DOI. Это закономерно проявляется в результатах Event Data об упоминаниях публикаций в российских журналах.

Сравним два журнала: российский «Вавиловский журнал генетики и селекции» и американский «Journal of the American Osteopathic Association». Оба имеют уникальные префиксы DOI, поэтому запрос API Event Data выводит результаты только для этих журналов (обычно префикс одинаков для нескольких журналов одного издательства). На запрос по событиям, связанным с публикациями из российского журнала (префикс 10.18699) с начала 2020 года, получаем пять записей; все они относятся к добавлению ссылок на статьи из журнала «Вавиловский журнал генетики и селекции» в Википедии, причем англоязычной. Аналогичный запрос для американского журнала дает более 1000 записей о событиях для DOI с префиксом 10.7556. Ограничив период только январем 2020 года, получаем 583 записи, фиксирующие упоминания в Твиттере (178), на новостных платформах (6) и в Википедии (399) — в 9 реализациях на разных языках. Разница очевидна.

В то же время для журнала «Успехи физических наук», давно и широко использующего DOI как для русскоязычной, так и для англоязычной версий [10], аналогичный запрос к API по префиксу 10.3367 только за период с 1 по 31 января 2020 года возвращает более тысячи записей; абсолютное большинство из них также относится к ссылкам на статьи из журнала в Википедии: армянской, английской, немецкой и сербской. Заметим, что в работе [8] показано, что CrossRef Event Data лучше всего отслеживает события именно в Википедии.

4.3. Определение статуса свободного доступа к публикациям с помощью сервиса Unpaywall

Сервис Unpaywall (<http://unpaywall.org/>) стал чрезвычайно популярен в последние годы. Он позволяет довольно точно определить для каждой публикации, имеющей DOI CrossRef, наличие свободно доступных версий как на сайтах издателей, так и в многочисленных репозиториях открытого доступа. Этот сервис интегрирован в авторитет-

ные базы данных Web of Science, Scopus, Dimensions и другие, обеспечивая возможность фильтрации результатов поиска по признаку доступности полных текстов. Тысячи библиотек по всему миру используют этот сервис для информирования пользователей о свободно доступных документах благодаря его интеграции в популярные link resolvers, такие как Primo, Summon, EBSCO Full Text Finder (последний — через API). Кроме того, бесплатное расширение Unpaywall, которое можно установить с адреса <http://unpaywall.org/products/extension> в браузеры Chrome и Firefox, определяет доступность документа на открытой в браузере странице, если находит на ней DOI CrossRef этого документа. Если статус доступности определен, на странице появляется значок «замочка»: серый, если документ закрыт, и зеленый, если он свободно доступен. Можно настроить расширение, чтобы цвет «замочка» для свободного доступа был разным: золотым, зеленым или бронзовым в соответствии с типом ОД, как он определен в [12].

Работа Unpaywall опирается на постоянно обновляемую базу данных, формирующуюся индексированием множества источников, перечень которых приведен на странице <http://unpaywall.org/sources>. Основа индексирования — база CrossRef; DOI является ключевым элементом записей в базе, к нему добавляется информация о найденных в других источниках открытых версиях документа с данным DOI и основаниях (evidence) для определения статуса документа как свободно доступного. Основаниями могут служить лицензия в метаданных CrossRef, лицензия на странице документа, факт принадлежности журналу открытого доступа или репозиторию открытого доступа, элемент метаданных репозитория в OAI-PMN, факт доступности файла pdf на странице публикации и другие. Информацию из базы данных CrossRef можно извлечь через API (<https://unpaywall.org/products/api>) или через онлайн-форму запроса с перечнем DOI (<https://unpaywall.org/products/simple-query-tool>)³. В работе [11] приведены количественные оценки использования разных типов оснований (evidence) в базе Unpaywall. На их основе версиям документа присваивается статус золотого, зеленого или бронзового открытого доступа — ОД (см. об этом в [12]).

К базе данных Unpaywall обращаются программы, регулярно проверяющие статус доступности документов для Web of Science и Scopus. Однако в пред-

³ Следует иметь в виду, что данные расширения Unpaywall, которое обращается к базе Unpaywall для определения статуса доступности документа, с конкретным DOI могут не совпадать с данными, полученными из базы по запросу об этом DOI. Возможная причина — в том, что результат анализа непосредственно на веб-странице документа еще не занесен в базу, которая обновляется с определенной периодичностью.

ставлении результатов обработки в этих базах есть разница: Scopus учитывает результаты без изменений, а Web of Science дополнительно присваивает статус золотого ОД всем записям на статьи из журналов, входящих в Directory of Open Access Journals (DOAJ), независимо от результатов работы Unpaywall, в частности в случае отсутствия DOI в статьях. Вследствие этого статьи из журнала, зарегистрированного в DOAJ и входящего одновременно в Web of Science и Scopus, могут получить разный статус доступности в этих базах.

Разумеется, результаты Unpaywall не являются безупречными. В [12] разработчики оценили полноту и точность работы сервиса как 77% и 97% соответственно; алгоритм постоянно совершенствуется с целью улучшения этих характеристик. Важно понимать, как работает сервис, чтобы выявить типовые ситуации, в которых определение статуса доступности документа становится затруднительным или невозможным. Это тем более важно в связи с тем, что на данных Unpaywall в базах Web of Science и Scopus делается оценка уровня открытости публикаций отдельных тематических направлений и различных стран [12, 13]. Можно выделить несколько типовых ситуаций, чтобы рекомендовать пути решения проблем для российских публикаций, авторы и издатели которых недостаточно учитывают роль CrossRef DOI для определения статуса ОД. Между тем миллионы пользователей во всем мире, ведущих поиск в авторитетных базах данных и библиотечных сервисах, которые используют link resolver, полагаются на Unpaywall и охотно обращаются к полным текстам документов, отмеченных как свободно доступные, тем самым увеличивая их видимость и вероятность цитирования [12, 14].

■ Отсутствие DOI

Пользователи, полагающиеся на работу расширения Unpaywall, часто недоумевают, почему оно не показывает зеленый статус документов на сайтах российских журналов, которые декларируют свободный доступ к своему контенту и действительно его предоставляют. Причина этого явления — в том, что статьи таких журналов не имеют DOI, и сервису не на что опереться, чтобы найти информацию о них в своей базе. Эта банальная причина имеет место даже в авторитетных журналах, входящих в «старшие» индексы Web of Science: Science Citation Index Expanded (SCIE), Social Science Citation Index (SSCI) и Arts & Humanities Citation Index (AHCI). Например, российский журнал «Вопросы психоло-

гии», входящий в SSCI, предоставляет в открытый доступ выпуски за двадцать лет, с 1980 по 1999 год, на своем сайте (<http://www.voppsy.ru/frame25.htm>). В SSCI находим 2804 статьи журнала за этот период; все они не определяются как свободно доступные, поскольку не имеют DOI⁴. Отметим, что за этот период Web of Science определяет количество свободно доступных статей российских авторов в 5283. Если прибавить к этому количеству только «потерянные» статьи из «Вопросов психологии», количество возрастет более чем в полтора раза.

■ Отсутствие элементов метаданных в базе CrossRef, позволяющих надежно определить статус открытости документа

Наилучшим способом предоставления информации о доступности документа в метаданных CrossRef является ссылка на лицензии Creative Commons, то есть включение в метаданные элемента License с значением, обозначающим соответствующий тип лицензии, например «<http://creativecommons.org/licenses/by/4.0/>». Российские журналы часто применяют лицензии Creative Commons или лицензии, по сути, соответствующие лицензиям Creative Commons, но не указывают это в метаданных CrossRef.

Unpaywall не ограничивается анализом метаданных CrossRef, программа анализирует и страницу самого документа и может обнаружить там дополнительные данные для определения статуса ОД, например явно прописанную ссылку на тот же Creative Commons или доступный для выгрузки pdf-файл, что поможет присвоить документу статус ОД, но это происходит далеко не всегда.

■ Отсутствие элементов, с помощью которых Unpaywall определяет статус ОД, на страницах (URL) документов

Как было сказано выше, Unpaywall анализирует страницу документа для определения статуса его доступности. При работе расширения Unpaywall программе необходимо найти на странице документа его DOI. Как определить, что DOI (которых на странице может быть несколько) является идентификатором данного документа? Самый надежный способ — включить эту информацию в метаданные на веб-странице в той структуре метаданных, которую использует платформа, например `<meta name="citation_doi" content="10.xxxx.xxxx">`. Это обеспечит корректное присвоение статуса открытого доступа и повысит комфортность работы пользователей. Так, обсуждение автором проблем работы расширения Unpaywall

⁴ Все количественные данные по Web of Science получены на массиве документов с 1975 года по настоящее время в последних числах июля 2020 года.

с командой разработчиков Научной электронной библиотеки завершилось тем, что они включили DOI публикаций в метаданные html-страниц — и теперь Unpaywall корректно показывает возможность свободного доступа для тысяч публикаций на популярной платформе elibrary.ru.

На рисунках 2 и 3 приведен пример статьи из журнала «Библиосфера», представленной на двух платформах: Elpub (<https://www.bibliosphere.ru/jour>) и платформе журналов Томского национального исследовательского университета (<http://journals.tsu.ru/bibliosfera/>). В первом случае Unpaywall находит DOI в метаданных, обращается к своей базе, определяет статус ОД статьи и демонстрирует значок зеленого замочка. Этот результат закономерен: платформа Elpub разработана с учетом современных требований к издательским сервисам, в том числе к метаданным, и отлажена на материале около 400 научных журналов, представленных сейчас на платформе. Во втором случае

Unpaywall не может определить статуса, поскольку DOI представлен на странице, но не в метаданных. Такая же ситуация имеет место для всех журналов на второй платформе⁵.

Если журнал предоставляет все статьи или часть их в свободном доступе, размещая на странице описания статьи ссылку на свободно доступный pdf-файл, Unpaywall может присвоить ей статус открытого доступа на основании этого факта. Но программе также нужна надежная информация о том, что эта ссылка — на pdf данной статьи, а не на упомянутые в тексте другие документы. И здесь хорошим решением снова является включение ссылки в метаданные, например так: `<meta name="citation_pdf_url" content=URL файла>`.

■ Ошибки в DOI, присвоенных статьям

Никто не застрахован от ошибок при регистрации и последующем использовании CrossRef DOI⁶. Это могут быть банальные опечатки, но могут быть более серьезные проблемы, проявляющи-

Рис. 2. Статья из журнала «Библиосфера» на платформе Elpub

Fig. 2. Article from “Bibliosphere” journal on Elpub platform

⁵ Ситуация зафиксирована по состоянию на 11.08.2020.

⁶ Распространенные типы ошибок при регистрации DOI в CrossRef проанализированы командой elpub.ru в https://elpub.ru/images/files/DOI_mistakes.pdf

Библиосфера
Библиосфера

ГЛАВНАЯ главная страница | **РЕДАКЦИЯ** редколлегия и редсовет | **АРХИВ** архив выпусков | **НОВОСТИ** новости и объявления | **ЭТИКА** Издательская этика | **АВТОРАМ** полезная информация

Цифровизация процессов обслуживания в библиотеках - это уже реальность | Библиосфера. 2019. № 2. DOI: 10.20913/1815-3186-2019-2-19-25

Поиск

Главная » Архив » Цифровизация процессов обслуживания в библиотеках - это уже реальность | Библиосфера. 2019. № 2. DOI: 10.20913/1815-3186-2019-2-19-25

Цифровизация процессов обслуживания в библиотеках - это уже реальность

Рассматриваются теоретические вопросы цифровизации в библиотеках, характеризуются современные формы и методы предоставления библиотечно-информационных услуг, способы коммуникации библиотек и пользователей. Раскрываются качественно новые возможности библиотечно-информационного обслуживания, обусловленные распространением цифровых технологий и робототехники. Примеры успешной реализации проектов цифровизации библиотечно-информационных услуг демонстрируют готовность библиотек к переходу на новый уровень развития. Автор приходит к выводу, что несмотря на то, что цифровизация существенно повышает оперативность предоставления услуг, положительно влияет на их качество и, следовательно, привлекает в библиотеки новых пользователей, библиотечно-информационное обслуживание включает также процессы, которые невозможны без участия человека, обладающего такими качествами, как способность к антиципации, эмоциональному переживанию, интуитивному прогнозированию. Необходимо осознавать и профессиональные потребности при использовании цифровых технологий в библиотечно-информационном обслуживании.

Скачать электронную версию публикации

Загружен, раз: 261

Рис. 3. Статья из журнала «Библиосфера» на журнальной платформе Томского университета

Fig. 3. Article from “Bibliosphere” journal on Tomsk University journal platform

еся не на единичных документах, а на их группах. Такие ошибки приводят к неприятным последствиям: ссылки на них во множестве других документов не будут работать. Обнаружить их трудно, но в ряде случаев может помочь именно анализ результатов работы Unpaywall. Обратимся за примерами к уже упоминавшейся журнальной платформе Томского университета. Она содержит несколько журналов, входящих в Web of Science, один из них — «Вестник Томского государственного университета. Математика и механика». Журнал зарегистрирован в ROAD — Directory of Open Access scholarly Resources. Его статьи свободно доступны, файлы pdf можно скачать с их страниц. В Web of Science находим 350 статей из этого журнала, из которых 342 имеют DOI, для 306 обозначен статус открытого доступа⁷. При формировании базы Unpaywall этот статус присвоен на основании наличия свободно доступных файлов pdf на страницах, адреса которых взяты из базы CrossRef. Статьи, которым Unpaywall

не присвоил статус свободного доступа, нужно проверить и определить причину. В частности, все статьи 3-го выпуска за 2016 год не опознаются CrossRef, то есть не находятся в базе данных; возможно, процедура регистрации не была завершена корректно.

5. Заключение. Выводы

DOI CrossRef играет все более важную роль в современной научной коммуникации. За два десятилетия своего существования в CrossRef создана более чем стомиллионная база описаний научных публикаций. На ее основе развиваются многообразные функции, позволяющие анализировать документальные потоки, отслеживать традиционное цитирование и новые типы упоминаний публикаций в Интернете. На ее основе развивается сервис Unpaywall, используемый для определения доступности полных текстов публикаций, к которым можно перейти от результатов

⁷ Данные поиска на 11.08.2020.

поиска в авторитетных базах данных и поисковых системах в тысячах библиотек мира. Знакомство с новыми функциями DOI CrossRef, понимание механизма их работы необходимо для эффективного использования возможностей, которые они дают авторам и издателям научных публикаций.

Учитывая эти тенденции, можно сформулировать ряд рекомендаций для авторов и издателей:

- присвоение и регистрация DOI CrossRef становится необходимым элементом введения публикации в научный оборот;
- метаданные публикации при регистрации DOI в CrossRef должны быть детальными и формализованными (в тех случаях, когда есть рекомендации на этот счет, например при вводе данных о лицензии);
- включение DOI публикаций должно стать необходимым элементом не только в списках литературы, но и при упоминании публикаций в социальных сетях, Википедии и других сетевых сообществах;
- журналам, практикующим открытый доступ ко всему контенту, следует ориентироваться на лицензии Creative Commons; регистрация таких журналов в Directory of Open Access Journals повысит видимость их публикаций как свободно доступных;
- на веб-страницах журнальных публикаций необходимо отражать в метаданных те элементы, которые помогают формально определить статус публикации как свободно доступной.

СПИСОК ЛИТЕРАТУРЫ

1. The Formation of CrossRef: a short history. URL: <https://www.crossref.org/pdfs/CrossRef10Years.pdf> (дата обращения: 05.08.2020).
2. Pentz E. CrossRef: The missing link. *College & research libraries news*. 2019;62(2):206–209.
3. Multiple resolution guide. URL: <https://www.crossref.org/get-started/multiple-resolution/> (дата обращения: 05.08.2020).
4. Walker J. CrossRef and SFX: complementary linking services for libraries. *New Library World*. 2002;103(3):83–89. <https://doi.org/10.1108/03074800210422296>
5. Crossref Fact File: 2018–2019 Annual Report. URL: <https://www.crossref.org/pdfs/annual-report-factfile-2018-19.pdf> (дата обращения: 05.08.2020).
6. Hendricks G., Tkaczyk D., Lin J., Feeney P. Crossref: The sustainable source of community-owned scholarly metadata. *Quantitative Science Studies*. 2020;1(1):414–427. https://doi.org/10.1162/qss_a_00022
7. Zahedi Z., Costas R. General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PloS One*. 2018 May 17;13(5):e0197326. <https://doi.org/10.1371/journal.pone.0197326>
8. Ortega J.L. Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event Data. *Scientometrics*. 2018;116(3):2123–2138. <https://doi.org/10.1007/s11192-018-2838-z>
9. Piwowar H., Priem J., Orr R. The Future of OA: A large-scale analysis projecting Open Access publication and readership. *BioRxiv*. 2019 Jan 1:795310. <https://doi.org/10.1101/795310>
10. Аксентьева М.С., Кириллова О.В., Москалева О.В. К вопросу цитирования в Web of Science и Scopus статей из российских журналов, имеющих переводные версии. *Научная периодика: проблемы и решения*. 2013;4(16).
11. Наупка Н. Exploring the Open Access Evidence Base in Unpaywall with Python. URL: https://subugoe.github.io/scholcomm_analytics/posts/unpaywall_python/ (дата обращения: 05.08.2020).
12. Пивовар Х., Прим Д., Ларивьер В., Алперин Х.П., Маттиас Л., Норландер Б. и др. Открытый доступ сегодня: широкомасштабный анализ распространенности и влияния статей открытого доступа. *Наука и научная информация*. 2019;2(4):228–247. <https://doi.org/10.24108/2658-3143-2019-2-4-228-247>
13. Богоров В.Г. Публикации открытого доступа в Web of Science: география революции. URL: <https://conf.neicon.ru/materials/63-Overseas2019/20190925-01-Bogorov.pptx> (дата обращения: 05.08.2020).
14. Archambault É., Côté G., Struck B., Voorons M. *Research impact of paywalled versus open access papers*. 2016.

REFERENCES

1. The Formation of CrossRef: a short history. Available at: <https://www.crossref.org/pdfs/CrossRef10Years.pdf> (accessed 5 August 2020).
2. Pentz E. CrossRef: The missing link. *College & research libraries news*. 2019;62(2):206–209.

3. Multiple resolution guide. Available at: <https://www.crossref.org/get-started/multiple-resolution/> (accessed 5 August 2020).
4. Walker J. CrossRef and SFX: complementary linking services for libraries. *New Library World*. 2002;103(3):83–89. <https://doi.org/10.1108/03074800210422296>
5. Crossref Fact File: 2018–2019 Annual Report. Available at: <https://www.crossref.org/pdfs/annual-report-factfile-2018-19.pdf> (accessed 5 August 2020).
6. Hendricks G., Tkaczyk D., Lin J., Feeney P. Crossref: The sustainable source of community-owned scholarly metadata. *Quantitative Science Studies*. 2020;1(1):414–427. https://doi.org/10.1162/qss_a_00022
7. Zahedi Z., Costas R. General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PloS One*. 2018 May 17;13(5):e0197326. <https://doi.org/10.1371/journal.pone.0197326>
8. Ortega J.L. Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event Data. *Scientometrics*. 2018;116(3):2123–2138. <https://doi.org/10.1007/s11192-018-2838-z>
9. Piwowar H., Priem J., Orr R. The Future of OA: A large-scale analysis projecting Open Access publication and readership. *BioRxiv*. 2019 Jan 1:795310. <https://doi.org/10.1101/795310>
10. Aksentyeva M.S., Kirillova O.V., Moskaleva O.V. On the issue of citing articles from Russian journals that have translated versions in Web of Science and Scopus. *Nauchnaya periodika: problemy i resheniya*. 2013;4(16) (In Russ.).
11. Haupka N. Exploring the Open Access Evidence Base in Unpaywall with Python. Available at: https://subugoe.github.io/scholcomm_analytics/posts/unpaywall_python/ (accessed 5 August 2020).
12. Pivovar Kh., Prim D., Lariviere V., Alperin H.P., Matthias L., Norlander B., et al. Open Access Today: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles. *Scholarly Research and Information*. 2019;2(4):228–247 (In Russ.). <https://doi.org/10.24108/2658-3143-2019-2-4-228-247>
13. Bogorov V.G. Open Access Publications in the Web of Science: The Geography of the Revolution (In Russ.). Available at: <https://conf.neicon.ru/materials/63-Overseas2019/20190925-01-Bogorov.pptx> (accessed 5 August 2020).
14. Archambault É., Côté G., Struck B., Voorons M. *Research impact of paywalled versus open access papers*. 2016.

ИНФОРМАЦИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Наталья Николаевна Литвинова, канд. филол. наук, эксперт Некоммерческого партнерства «Национальный электронно-информационный консорциум» (НЭИКОН); главный библиотекарь Российской государственной библиотеки;
nlit@neicon.ru
 ORCID: <https://orcid.org/0000-0002-9077-6288>

Natalia N. Litvinova, Cand. Sci. (Philol.), Expert of the Non-profit Partnership “National Electronic Information Consortium” (NEICON); Chief Librarian of the Russian State Library;
nlit@neicon.ru
 ORCID: <http://orcid.org/0000-0002-9077-6288>