

НАУЧНЫЕ КОММУНИКАЦИИ / SCIENTIFIC COMMUNICATIONS

Дискуссионная статья / Discussion paper

<https://doi.org/10.24069/SEP-24-20>

Идентификатор DOP: новый метод идентификации научных публикаций

Х. Мустафа  

Арабский сервер препринтов (ArabiXiv), г. Париж, Франция

 khaled.moustafa@arabxiv.org

Резюме. Эффективная система идентификации научных публикаций должна содержать такую важную информацию, как имя автора, название журнала, в котором опубликована статья, и дату ее публикации. Однако существующие постоянные идентификаторы (persistent identifiers, PIDs), как правило, не содержат таких данных, поскольку основная цель этих идентификаторов – связывать онлайн-публикации с их веб-источниками. В печатных версиях постоянные идентификаторы теряют свою функциональность, так как предназначены исключительно для поиска контента в интернете. Для решения этой проблемы был предложен новый метод идентификации и связывания научных публикаций – идентификатор DOP (date of publication – «дата публикации»). Он основан на четырех ключевых элементах: 1) имени первого автора, 2) названии журнала, 3) дате публикации (DOP), 4) времени публикации. Благодаря этим параметрам идентификатор DOP позволяет эффективно идентифицировать публикации как в цифровом, так и в печатном формате. Например, статья, опубликованная автором А в журнале J в дату D и время T, будет иметь идентификатор DOP следующего формата: DOP>AuthorA/JournalJ/D/T, где D обозначает дату публикации (год, месяц, день), а T – время публикации (час, минута, секунда). Данный подход имеет ряд преимуществ: DOP содержит основную информацию прямо в ссылке; уникальность идентификаторов достигается за счет точной даты и времени публикации; формат DOP совместим с привычными способами цитирования, использующими аналогичные переменные; даже если исходная ссылка будет перемещена, удалена или недоступна, идентификатор DOP сохранит все основные данные для идентификации публикации. Кроме того, идентификаторы DOP можно использовать для создания глобальной библиографической базы данных, которая обеспечит централизованную идентификацию и индексирование научных публикаций. Система идентификаторов DOP предлагает универсальный, надежный, прозрачный и удобный способ идентификации научных работ, обеспечивающий их доступность и корректную атрибуцию.

Ключевые слова: научная публикация, постоянный цифровой идентификатор, PID, цифровой идентификатор объекта, идентификатор DOI, идентификатор DOP, дата публикации, индексная база данных, идентификационная связь, открытая архивная база данных, инструмент с открытым исходным кодом, библиографическая база данных, библиографические индексы, идентификация рукописей

Для цитирования: Мустафа Х. Идентификатор DOP: новый метод идентификации научных публикаций. *Научный редактор и издатель*. 2024;9(2):221–226. <https://doi.org/10.24069/SEP-24-20>

DOP identifier: A new method of identification of scholarly publications

K. Moustafa  

The Arabic Preprint Server (ArabiXiv), Paris, France

 khaled.moustafa@arabxiv.org

Abstract. An effective identification system for scholarly publications should include essential identifying information, such as the author's name, publishing journal, and date of publication. However, current Persistent Identifiers (PIDs) often lack such important details, as their primary goal is to link online publications to their web sources. In print versions, these PIDs become ineffective since their main purpose is to locate online content only. To address these limitations, a new identification and linking system, called DOP (Date of Publication) is proposed, based on four key variables: 1) the first author's name, 2) the journal's name, 3) the date of publication (DOP), and 4) the time of publication. With such identifying features, DOP links can identify scholarly publications efficiently both in print and online settings. For instance, a paper published by author A in journal J on date D and time T can have a DOP identifier link as follows: DOP>AuthorA/JournalJ/D/T, where D refers to the date of publication (year, month, and day), and T refers to the time of publication (hour, minute, and second). This format provides several advantages: it offers key identifying information within the links themselves; the date and time make DOP identifiers specific among journals and between authors; it fits with common citation styles that typically rely on similar variables; and even if the original publication link turns out to be broken, moved, or no longer exists, informative data remain visible within the DOP link itself. Furthermore, a large bibliographic database can be established based on DOP identifiers for the identification and indexation of scientific publications in one and the same place. DOP identifiers can thus offer a global, robust, cohesive, clear and human readable way to identify scholarly publications from around the world.

Keywords: scholarly publication, persistent digital identifier, PID identifier, digital object identifier, DOI identifier, DOP identifier, date of publishing, date of publication, indexation database, identification linkage, open archiving database, open-source tool, bibliography database, bibliography indexes, manuscript identification

For citation: Moustafa K. DOP identifier: A new method of identification of scholarly publications. *Science Editor and Publisher*. 2024;9(2):221–226. <https://doi.org/10.24069/SEP-24-20>

Идентификатор DOP (дата публикации): новый подход к идентификации научных публикаций и созданию ссылок на них

Научные публикации – это результаты интеллектуальной деятельности, создаваемые авторами и публикуемые в научных журналах. Они характеризуются уникальными названиями и датами выхода. Библиографическая ссылка на такие работы, как правило, включает указание имени автора, названия журнала, даты публикации и заголовка статьи. Для обеспечения информативности, удобства восприятия и соответствия общепринятым способам цитирования надежная система идентификации должна учитывать все эти характеристики или как минимум их часть. Однако имеющиеся в настоящее время постоянные идентификаторы (persistent identifiers, PIDs)

часто используют сложные комбинации чисел и символов, что далеко не всегда является эффективным и удобным способом идентификации научных публикаций [1].

Чтобы решить эти проблемы и создать эффективный способ идентификации научных публикаций, предложен новый метод ссылок и идентификации под названием DOP (date of publication – «дата публикации»), основанный на использовании четырех ключевых характеристик любой научной работы:

- имени автора (author's name);
- названия журнала (journal's name);
- даты публикации (date of publication, DOP);
- времени публикации (time of publication).

В данной системе идентификации «имя автора» означает фамилию основного автора, а «название журнала» – полное наименование журнала

или общепринятое сокращение. «Дата публикации» состоит из года, месяца и дня, а «время публикации» – из часа, минуты и секунды.

Эти характеристики одновременно уникальны и универсальны, поскольку каждая научная публикация определяется такими элементами, как имя автора, платформа публикации (обычно журнал или веб-сайт), дата и точное время выхода. Заглавия статей, несмотря на их важность для идентификации, в данном случае исключены, так как они часто бывают слишком длинными. Такое решение делает идентификаторы DOP более лаконичными и удобными для восприятия. Их стандартизованный формат с четкой последовательностью элементов обеспечивает информативность, удобочитаемость и доступность для пользователей.

Например, идентификатор DOP может иметь следующий формат:

>DOP: FirstAuthorLastName/JournalName(OR) abbreviation/year-month-day/hour-minute-second.

Чтобы идентификаторы DOP оставались короткими, доменные имена (например, .org, .com, .net и др.), префикс www и заглавия публикаций (часто слишком длинные) можно опустить.

Статья, опубликованная автором А в журнале J 31 декабря 2024 г. в 23 часа 59 минут 59 секунд (один из самых длинных возможных форматов даты), может быть идентифицирована следующим образом:

>DOP: AuthorA/JournalJ/Y2024M12D31/
hh23mm59ss59,

где Y – год; M – месяц; D – день; hh – часы; mm – минуты; ss – секунды.

Еще один возможный формат:

DOP: FirstAuthorLastName/JournalName/
Date(Y-M-D)/Time(h-m-s).

Для этого формата идентификатор DOP будет выглядеть следующим образом:

>DOP: AuthorA/JournalJ/Y2024M12D31/
T23h59mm59s.

Кроме того, формат может комбинировать или объединять переменные даты и времени (без пробелов или других разделителей). В идентификаторе

>DOP: AuthorA/JournalJ/D20241231/T235959

«D20241231» обозначает дату 31.12.2024, а «T235959» обозначает время 23 часа 59 минут 59 секунд.

Идентификаторы DOP могут быть записаны еще короче с помощью одноразрядного форма-

та даты и времени без указания нулей там, где это возможно. Например, следующий идентификатор DOP «AuthorA/JournalJ/D2511/T567» может идентифицировать статью, опубликованную 1 января 2025 г. в 5 часов 6 минут 7 секунд автором А в журнале J, где «D2511» обозначает дату (2025 год, 1-й месяц, 1-й день), а «T567» обозначает время 5 часов 6 минут 7 секунд.

Использование четырех переменных – имени автора, названия журнала, даты и времени публикации – позволяет формировать различные комбинации идентификаторов, обеспечивая их гибкость и уникальность:

– DOP: FirstAuthorLastName/JournalName/uuuuMMdd/hhmmss,

– DOP: JournalName/FirstAuthorLastName/yymd/hhmmss,

– DOP: JournalName/FirstAuthorLastName/D/T,

– DOP: FirstAuthorLastName/D/T/JournalName.

Однако каждому журналу или издательству следует четко придерживаться единого формата идентификаторов DOP для обеспечения стандартизации и согласованности идентификации публикаций. Это означает, что журналы и издательства могут выбирать собственный формат идентификатора DOP, но внутри одного издания или издательства этот формат должен оставаться неизменным. Такой подход позволяет сохранить внутреннюю согласованность, а также гибкость в работе с различными публикациями. Идеальным вариантом, который обеспечил бы максимальную совместимость, стал бы единый формат идентификатора DOP для всех журналов и издательств, что способствовало бы беспрепятственной интеграции и единству научной экосистемы. Но если учесть все разнообразие стилей цитирования и форматов идентификации, которые используются в разных журналах, становится очевидным, что добиться такого уровня унификации будет непросто.

Преимущества системы идентификаторов DOP

У предложенной системы идентификации DOP имеются следующие преимущества:

1. Уникальность. Включение в идентификаторы DOP переменных даты и времени с точностью до секунды делает эти идентификаторы уникальными, практически исключая вероятность дублирования. Маловероятно, что один автор опубликует в конкретном журнале в точно тот же момент времени более одной статьи. Использование минут и секунд в сочетании с датой обеспечивает практически бесконечные возможности для

создания уникальных идентификаторов DOP, гарантирует ясность и однозначность идентификаторов и помогает избежать путаницы и ошибок. Привязка к дате и времени – это эффективный метод как для идентификации, так и для архивирования.

2. Информативность. Использование имени автора, названия журнала, даты и времени публикации позволяет идентификаторам DOP сразу передавать контекст цитируемой работы. Хотя такая идентификация может удлинить URL, в данной системе приоритет отдается точности, а не краткости ссылок. Если рукописи имеют несколько версий (например, препринты), идентификаторы DOP отражают хронологию их публикации, порядок и делают каждую версию легко узнаваемой. Это полезно и для публикаций на разных языках, так как указывающие на них ссылки на основе идентификаторов DOI могут затруднять их идентификацию [2]. Идентификаторы DOP, в которых указываются дата и время, позволяют легко отличить оригинальную версию от переводной, поскольку оригинал всегда публикуется раньше перевода.

3. Универсальность для онлайн- и офлайн-среды. Ранее журналы не имели, а в некоторых случаях до сих пор не имеют идентификаторов для печатных или онлайн-публикаций [3]. Идентификаторы DOP устраняют разрыв между печатными и онлайн-изданиями и обеспечивают доступ к основной идентификационной информации независимо от формата публикации или наличия интернет-соединения. Современные цифровые постоянные идентификаторы (PIDs) не подходят для использования в печатных изданиях. Так как основная задача PIDs – локализация онлайн-контента, они становятся бесполезными для физических копий публикаций. Идентификаторы DOP решают эту проблему, так как они изначально содержат ключевую информацию – имя автора, журнал и дату публикации – и не требуют перехода по ссылке для получения этих данных. Эти «встроенные» данные обеспечивают точность и надежность идентификации как в онлайн-, так и в офлайн-среде.

4. Согласованность со способами цитирования. Разные форматы цитирования обычно включают такие элементы, как имя автора, название журнала, название публикации и дата ее выхода, при этом порядок и стиль их представления могут варьироваться. Идентификаторы DOP изначально содержат три ключевых элемента: имя автора, название журнала и дату публикации. Этот подход обеспечивает соответствие основ-

ным стилям цитирования, снижает вероятность ошибок и облегчает корректную атрибуцию, одновременно гармонизируя формат цитирования с представленными ссылками.

5. Эффективный поиск литературы. Информативность идентификаторов DOP позволяет читателям быстро оценить актуальность публикаций, всего лишь взглянув на указанную в ссылке дату. Это помогает сэкономить время при поиске литературы и сосредоточиться на публикациях, относящихся к нужным временным рамкам. Например, благодаря дате, указанной в идентификаторе DOP, читатель может легко понять, является ли публикация новой или устаревшей, и ему не приходится для этого скачивать или открывать файлы. Такой подход значительно повышает эффективность поиска и облегчает доступ к интересующим материалам, как самым свежим, так и более ранним.

6. Возможность создания единой базы данных. Идентификаторы DOP могут стать основой для создания глобальной библиографической базы данных научных публикаций, тем самым способствуя открытому равноправному распространению научной информации и обеспечивая к ней доступ широкого круга издателей и журналов по всему миру. На основе идентификаторов DOP можно создать библиотеку, позволяющую индексировать и находить библиографическую информацию по научной литературе, опубликованной в журналах и издательствах разных стран, – все в одном месте. Этот уникальный хаб DOP может предоставлять доступ к полным текстам статей, если они доступны, или при их отсутствии предлагать краткую аннотацию с основной информацией для цитирования.

7. Простота внедрения. Система идентификаторов DOP может быть внедрена достаточно просто с использованием плагинов или программного кода, который объединяет четыре ключевых параметра (имя автора, название журнала, дату и время публикации) в доступные ссылки по аналогии с другими цифровыми идентификаторами. Автоматизация этого процесса позволит обеспечить точность и единообразие идентификаторов DOP на разных платформах, значительно упростив их создание и использование в научной среде.

Внедрение системы идентификаторов DOP может стать частью более широкой стратегии, направленной на разработку разнообразных решений для удовлетворения потребностей издателей, журналов и научных сообществ. Такой информативный, универсальный и удобный инструмент,

как DOP, способен увеличить функциональность академических баз данных и библиографических индексов, повысить удобство использования и сделать поиск литературы и работу со ссылками более эффективными.

Ограничения и сложности использования идентификаторов DOP

1. Разнообразие форматов даты и времени. Различные форматы даты и времени в разных часовых поясах могут усложнить стандартизацию идентификаторов DOP. Однако эту проблему можно решить следующими способами:

- использовать единый формат или стандартизированное время, например универсальное координированное время (UTC). Это можно реализовать путем добавления формата UTC в качестве префикса или суффикса аналогично префиксу «10.12345» в DOI, используемому всеми издателями и журналами;

- использовать трехбуквенные сокращения для месяцев (например, Jan., Feb., Mar.);

- использовать буквы для обозначения элементов даты публикации (Y для года, M для месяца, D для дня);

- использовать 24-часовой формат для обозначения времени публикации (hh для часов, mm для минут, ss для секунд), чтобы избежать путаницы со временем до/после полудня (AM/PM).

Пример:

DOP: AuthorA/JournalJ/Y2025M12D31/
hh11mm15ss45,

где AuthorA – автор; JournalJ – журнал; 2025 – год публикации; 12 – декабрь (месяц); 31 – день месяца; hh11mm15ss45 – время публикации в 24-часовом формате (11:15:45 утра).

2. Изменения в названиях журналов и именах авторов. Названия журналов и имена авторов могут меняться со временем, но такие изменения не должны влиять на работу идентификаторов DOP. Идентификаторы должны указывать на сами публикации и их местоположение на момент выхода, служа своего рода «снимком» оригинального состояния публикации. Все изменения должны отражаться на сайтах журналов, а идентификаторы DOP – перенаправлять на новое местоположение публикации, если это необходимо.

3. Имена и специальные символы. Включение в идентификатор фамилии автора не всегда может быть оправдано из-за культурных различий и особенностей структуры имен. Однако использование фамилий в библиографических ссылках на научную литературу является обще-

принятым стандартом и в контексте идентификаторов DOP не должно считаться нарушением культурных норм.

В некоторых языках имена могут быть написаны с использованием специальных символов или нелатинскими буквами, что также может повлиять на стандартизацию DOP. Однако такие современные веб-технологии, как UTF (Universal character set transformation format), то есть система кодирования, работающая по стандарту Unicode, успешно справляются с этими проблемами. Чтобы избежать каких-либо ошибок, издатели должны просто копировать такие имена из рукописей или форм подачи и вставлять их в ссылку идентификатора DOP.

4. Стоимость и доступность. Создание системы идентификаторов DOP, как и любого другого продукта или услуги, требует финансовых вложений, опытных программистов и грамотного маркетинга. Тем не менее ее простота и логичность может помочь снизить затраты на ее разработку. Совместные усилия финансирующих организаций, издателей, библиотек, разработчиков и сообществ, работающих в области открытой науки, могут привести к созданию надежного, доступного, а возможно, и полностью бесплатного решения. Разработка системы идентификаторов DOP как некоммерческого проекта, открытого для научного сообщества, позволит избежать конфликтов интересов, связанных с финансовыми вопросами, а также повысит ее прозрачность и доступность.

Несмотря на перечисленные ограничения, система идентификации DOP обладает значительными преимуществами в плане информативности и практической ценности. Открытая и прозрачная структура этой системы упрощает доступ к научной информации и ее использование, предлагая тем самым более эффективное решение по сравнению с существующими системами.

Заключение

Предложенная система идентификации научных публикаций DOP обладает значительными преимуществами по сравнению с существующими методами. Идентификаторы DOP, содержащие ключевые элементы – имя автора, название журнала, дату и время публикации, – обеспечивают понятный, удобочитаемый и соответствующий требованиям цитирования способ идентификации. В отличие от других способов идентификации (PIDs), использующих сложные символы и числа, идентификаторы DOP точнее соответствуют основным принципам цитирования, снижая вероятность ошибок.

Точные временные метки упрощают сравнение различных версий публикаций, создавая четкий хронологический порядок. Это особенно важно для установления приоритета открытий и защиты авторских прав. Идентификаторы DOR универсальны и подходят для всех типов научных работ как в печатном, так и в цифровом формате, включая учебные пособия, материалы конференций и технические отчеты.

Кроме того, идентификаторы DOR упрощают архивирование и доступ к публикациям, устраняя зависимость от внешних ссылок или систем доступа. Идентификаторы DOR также могут спо-

собствовать созданию единой системы идентификации в рамках отдельного журнала или издательства, обеспечивая более надежный, прозрачный и удобный способ работы с научной информацией. С помощью идентификаторов DOR, предоставленных различными журналами и издательствами, можно создать глобальную библиографическую базу данных, содержащую основные данные для индексирования и цитирования и предоставляющую доступ к полным текстам статей.

Перевод: Бюро переводов «TextTranslate»

КОНФЛИКТ ИНТЕРЕСОВ

Автор заявляет об отсутствии у него конфликта интересов. Финансирование работы отсутствовало.

CONFLICT OF INTERESTS

The author declares no conflict of interests. There was no funding for the work.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Moustafa K. DOI illusion. *Revista Clínica Española*. 2024;224(5):334–335. <https://doi.org/10.1016/j.rceng.2024.04.002>
2. Бирюков А.А., Скалабан А.В. Технические аспекты публикации на нескольких языках – как правильно связать цифровые идентификаторы объектов (DOI). *Научный редактор и издатель*. 2020;5(1):29–39. <https://doi.org/10.24069/2542-0267-2020-1-29-39>
Birukou A.A., Skalaban A.V. Technical aspects of publication in several languages – how to link digital object identifiers (DOIs). *Science Editor and Publisher*. 2020;5(1):29–39. (In Russ.) <https://doi.org/10.24069/2542-0267-2020-1-29-39>
3. Moustafa K. Fake journals: Not always valid ways to distinguish them. *Science and Engineering Ethics*. 2015;21(5):1391–1392. <https://doi.org/10.1007/s11948-014-9608-y>

ИНФОРМАЦИЯ ОБ АВТОРЕ

Халед Мустафа – PhD, основатель и редактор Арабского сервера препринтов (ArabiXiv), Париж, Франция; <https://orcid.org/0000-0002-0819-9064>; e-mail: khaled.moustafa@arabixiv.org

INFORMATION ABOUT THE AUTHOR

Khaled Moustafa – PhD, Founder and Editor of the Arabic Preprint Server (ArabiXiv), Paris, France; <https://orcid.org/0000-0002-0819-9064>; e-mail: khaled.moustafa@arabixiv.org

Поступила в редакцию / Received 06.06.2024

Поступила после рецензирования / Revised 12.11.2024

Принята к публикации / Accepted 04.12.2024