
И.А. Тихонкова, канд. биол. наук

Старш. науч. сотр. Института молекулярной биологии
и генетики НАН Украины



DOI (DIGITAL OBJECT IDENTIFIER) – ОБЯЗАТЕЛЬНЫЙ ЭЛЕМЕНТ СОВРЕМЕННОГО НАУЧНОГО ИЗДАНИЯ

Появление IT технологий и Интернета коренным образом изменили не только издательский процесс, но и методологию поиска научной информации. Несмотря на наличие у большинства научных изданий собственных веб-сайтов часто возникают сложности с поиском конкретной публикации. В конце 1990-х годов было предложено использовать *doi (digital object identifier)* идентификатор цифрового объекта для создания постоянных гиперссылок, которые бы работали (приводили нас к искомым веб-страницам) даже при изменении структуры сайта. За эти годы данный индекс стал широко использоваться как научными изданиями, так и учеными, а дополнительные возможности, которые предлагаются фирмой *CrossRef* становятся обязательными для журналов высокого уровня. К сожалению, на сегодняшний день *doi* имеют лишь два украинских журнала, еще у нескольких он предоставляется для переводных версий журналов, переиздающихся за рубежом. Целью данной статьи является ознакомление с возможностями *doi* и других приложений *CrossRef* и практические рекомендации для журналов как получить данный индекс.

Как известно, каждое исследование начинается с детальной проработки литературы по интересующей тематике. В "доинтернетовскую эпоху" в библиотечный день ученый отправлялся в библиотеку, просматривал каталоги новых поступлений, в заинтересовавшей его статье обрабатывал списки использованной литературы, после чего начинал поиски первоисточников, которые, зачастую, были достаточно трудоемкими, а ожидание журнала с нужной статьей занимало дни или недели. Интернет коренным образом изменил поиск научной информации, но по-прежнему существуют некие подводные камни, о которых необходимо знать, чтобы успешно осуществлять научные изыскания. Кажется, что в Интернете за считанные минуты можно найти информацию по любой тематике, тем более, что большая часть



уважающих себя и читателей изданий постепенно оцифровывают и размещают на собственных веб-страницах не только текущие, но и архивные номера. Следовательно, на веб-сайте журнала в открытом доступе есть резюме (заголовок статьи, авторы, места выполнения работы, ключевые слова) и выходные данные статьи (название журнала, год издания, том, номер, страницы). Это позволяет роботам поисковых систем, при корректной структуре сайта журнала, находить искомую статью по названию практически мгновенно. Иначе выглядят результаты поиска по ключевым словам. Любые поисковые системы "Гугл", "Яндекс" и др. выдают большое количество статей, часть из которых может быть очень далека от научной периодики. Поэтому, как правило, ученые для поиска научной информации пользуются тематическими базами данных, где индексирование конкретного журнала, в основном, определяется комиссией специалистов данной отрасли науки. Некоторые базы являются открытыми: *PubMed/MedLine*, *elibrary.ru*, *DOAJ* и др., пользование же другими – платное: *Scopus*. Существуют также агрегаторы/библиотеки научных изданий *EBSCO*, *HINARI* и т. д., представляющие доступ к коллекциям журналов, данные базы также имеют удобные поисковые интерфейсы и ученый при наличии, например, институтской подписки на данную коллекцию может не только найти интересующую статью, но и сразу прочитать ее. Поэтому каждый научный журнал должен быть представлен (или стремиться попасть) в указанные или другие, профильные его области науки, базы. Наукометрические базы данных *Scopus (SJR)*, *Web of Science*, *elibrary.ru* анализируют публикационную активность издания, присваивая изданиям различные индексы (импакт-факторы, индекс Хирша и т. д.), позволяющие при продуманном подходе ранжировать журналы. В тематических базах данных типа *Scopus*, *PubMed*, *elibrary.ru* присутствуют и доступны читателями выходные данные статьи, резюме, и довольно часто список литературы и ссылка на сайт издательства. В зависимости от базы и заключенного договора эта информация или подается изданием, или добавляется сотрудниками базы самостоятельно на платной или бесплатной основе из электронной или печатной версии журнала. И, кажется, если издание реферировается такой базой, и с ее сайта есть гиперссылка на сайт издания, то поток читателей будет расти. Но, работая в Интернете и переходя по гиперссылкам, каждый из нас хотя бы раз попадал на страницу ошибки 404 или *Not found* (рис. 1), которая информирует нас о том, что данной статьи не существует.

Если мы имеем в виду новостной сайт или Интернет-магазин, действительно, эту статью могли удалить, а в случае журналов причина, скорее всего, в том, что издатели изменили гиперссылку, связывающую данные страницы. Причинами этого могут быть: изменение структуры сайта журнала, переход на другую платформу издания, а также смена владельца издания, что бывает с иностранными журналами.

К сожалению, у издателя не всегда есть возможность обновить URL статьи (ее адрес в Интернете), особенно если сотрудники базы сами вносят ин-

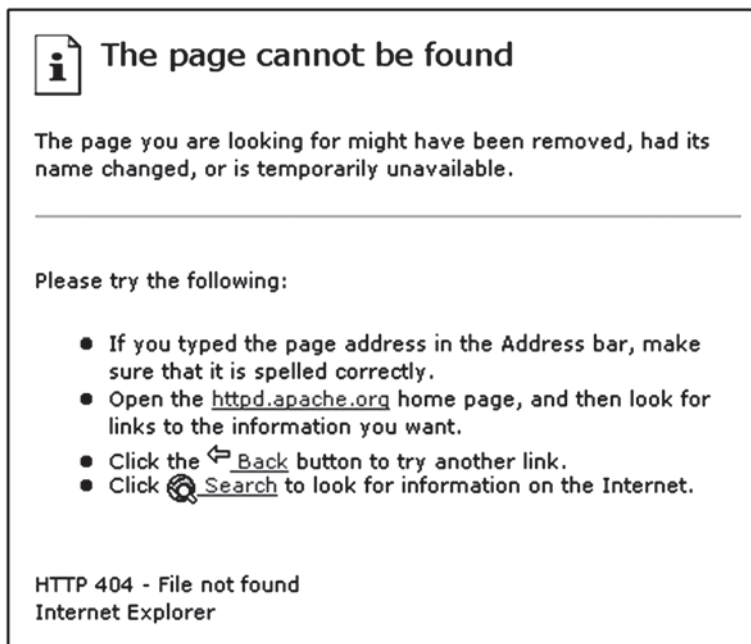


Рис. 1. Скриншот страницы ошибки 404

формацию о журнале или если это касается архивных номеров. И иногда с сайта *PubMed* вместо сайта журнала мы попадаем на "ошибку 404".

Кроме того, ссылки на отдельные статьи или номера Вашего журнала могут храниться в более узкоспециализированных базах данных, репозиториях или у читателя на компьютере, о чем вы можете не знать и не иметь возможности сообщить об изменении *URL*, и в таких случаях замена на корректную ссылку не возможна. И попытка перехода по устаревшей ссылке на публикацию в Вашем издании также может закончиться страничкой ошибки 404.

Для создания механизма постоянно действующих ссылок в конце 90^х гг. прошлого столетия было предложено использование индекса *doi* (*digital object identifier*) – идентификатор цифрового объекта. Необходимо отметить, что созданная в 1998 *International DOI Foundation* осуществляет общее руководство, внедряет стандарты, а с издателями непосредственно сотрудничает сеть регистрационных агентств, каждое из которых специализируется на определенном содержании. Регистрационное агентство *CrossRef*, работающее с 2000 года, присваивает данный индекс научным публикациям, книгам, материалам конференций. *Japan Link Center (JaLC)* регистрирует *doi* материалов на японском языке, *Airiti, Inc.* – на китайском; *R.R. Bowker* – книги и сериальные издания; *EIDR* – различные фильмы, телепередачи, аудио и т. д.



Рис. 2. Схема doi

Каждый научный сотрудник не раз видел данный индекс в статьях, на сайтах журналов, но, возможно, не каждый знает, для чего он нужен и как его применять.

Doi это уникальная алфавитно-цифровая строка, присвоенная цифровому объекту (статье, главе книги). В системе *CrossRef* каждый *doi* связан с набором базовых метаданных и *URL*-ссылке на полный текст, т. е. однозначно идентифицирует элемент содержания и обеспечивает постоянную ссылку на его местоположение в Интернете. Данный индекс построен по схеме, представленной на рис. 2: *doi* директория, префикс, суффикс. *Doi* директория <http://dx.doi.org> может не указываться в печатной версии статьи. Префикс – в данном примере 10.7124 (цифра 10, точка и четыре или пять цифр) присваивается регистрационным агентством издателю и указывает, кто издал данную статью, но не показывает, кому в данный момент принадлежит издание, т. е. при смене владельца издания индекс *doi* не меняется. Суффикс (bc.000027), присваивается издателем конкретной публикации или ее части. Требования к суффиксу: уникальность в пределах префикса (у одного издателя). Формат суффикса определяет издатель, можно использовать буквы латинского алфавита, цифры и определенные символы.

На схеме рис. 3 представлен упрощенный процесс издания статьи. После того, как статьи приняты в печать и подготовлена рукопись (т. е. издателю известны полные выходные данные всех статей номера: название, фамилии авторов, места работы, год, том, номер, страницы), создается веб-страничка выпуска журнала и так называемые "посадочные" (*landing*) или "целевые страницы" (*target page*) каждой статьи. Соответственно, у каждой страницы появляется свой адрес *URL*, например <http://biopolymers.org.ua/doi/bc.000027>. На "целевой странице" должны быть: резюме, выходные данные и ссылка на полный текст статьи (важно, что доступ к полному тексту может быть как свободным, так и по подписке, т. е. данный индекс никак не ущемляет бизнес-интересы издательства).

Далее издатель присваивает суффикс каждой статье, префикс он получает при подписании контракта с *CrossRef*. Таким образом, у издателя есть триада: исходные данные статьи, *URL* статьи и индекс *doi* (префикс + суффикс). Данная информация подается в регистрационное агентство (*CrossRef*). В верстке статьи размещается *doi* (в нашем случае – *doi*: 10.7124/bc.000027, и статья отправляется в печать.

CrossRef (*DOI*) можно представить каталогом, в котором на одной "карточке статьи" указаны три типа данных: выходные данные статьи, *doi* и *URL*,

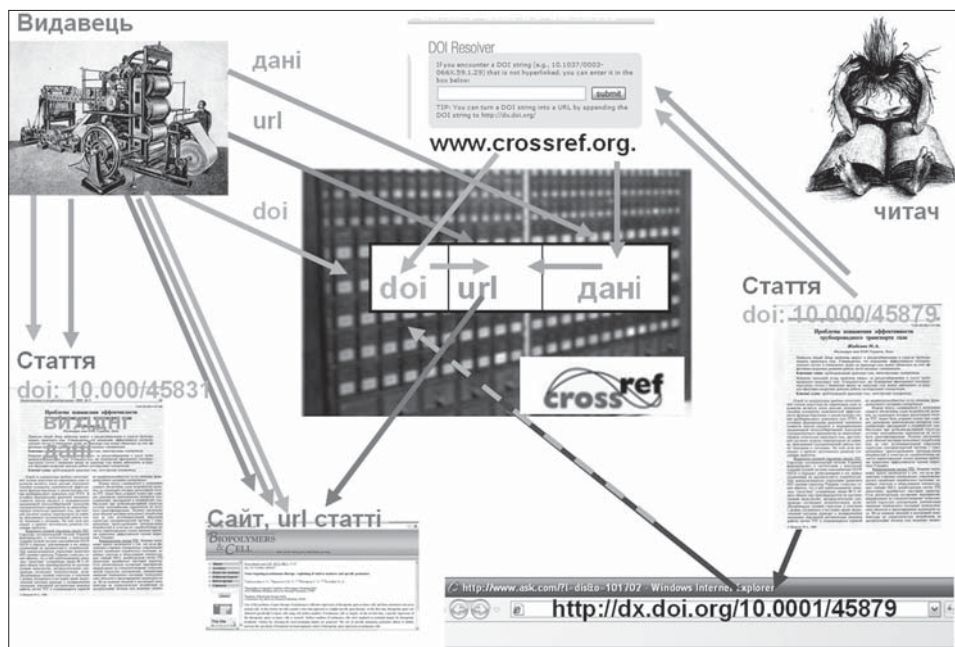


Рис. 3. Схема функционирования doi

два первых являются неизменными у данной публикации, а URL издатель может обновлять при необходимости. Важно подчеркнуть, что CrossRef – это не база полнотекстовых статей или поисковая система, а наличие индекса doi не является признаком "качественности" издания, так как любой издатель, подписав контракт с CrossRef и оплатив годовой членский взнос, получает префикс, имеет право создавать суффиксы и присваивать своим статьям doi. Размер годового взноса для каждого издательства определяется его годовым оборотом (от 275 долларов США). Стоимость doi для одной статьи текущего содержания (последние три года) составляет один доллар США, а для более ранних статей 15 центов – фиксированная цифра, не зависящая от масштабов издательства. Это разовая плата, т. е. даже при замене издателем URL доплата не требуется. Данные расценки могут позволить себе даже самые небольшие издательства, издающие один журнал. Читатели ничего не платят, пользуясь данным индексом. Как же читателю, зная индекс doi, найти статью? Открыв сайт DOI корпорации <http://www.doi.org> или CrossRef <http://www.crossref.org>, подставив doi в окно "Resolve a DOI Name", Вас перенаправят на "целевую страничку" данной статьи. Другой способ: если к doi индексу 10.7124/bc.000027 дописать doi директорию <http://dx.doi.org/> и подставить полученную ссылку <http://dx.doi.org/10.7124/bc.000027> (иногда doi сразу представлен в таком виде) в строку браузера, Вы также попадаете на "целевую страничку".



На сегодняшний день более 83 миллионов цифровых объектов имеют индекс *doi*, из которых свыше 60 миллионов определяют научные публикации. *Doi* можно присвоить книге или отдельной главе книги, что немаловажно, когда главы написаны разными авторами; материалам конференций; отчетам; диссертациям; журналу; тому; номеру; статье; определенным ее частям: схеме, рисунку, графику, таблице и т. д.

Иногда данный индекс присваивается статье, прошедшей рецензирование, но не готовой в окончательном печатном варианте так называемой допечатной публикации "*ahead of print*". В таком виде статья может находиться около года, размещенной на сайте издательства или базах данных, а цитировать ее можно лишь используя *doi*.

Для читателя ценность *doi* понятна. В любое время можно быстро найти статью, нет необходимости искать сайт журнала, в нем год, том, номер журнала, легко поделиться ссылкой с коллегами и т. д. А в чем необходимость наличия *doi* для журнала? Тем более, что именно на его плечи ложится оплата данного индекса.

Не все журналы представлены в *PubMed*, *Tompson Reuters* или *Scopus*, что не всегда свидетельствует о низком научном уровне. Возможно, он не подходит по профилю данной базы, но это не означает, что его не читают и не цитируют. Как привлечь новых читателей, узнать, кем процитирована данная статья, избежать плагиата, указать на наличие обновлений/уточнений вышедших статей? Все эти задачи/вопросы можно решить без *doi*, однако временные затраты в этом случае огромны. *CrossRef* помимо регистрации индекса *doi*, также предлагает ряд услуг издательствам. Например, *Cyted-by linking* – сервис, показывающий кем и где определенная статья процитирована. У некоторых журналов на сайте есть закладки "*This article was cited by*", по которой можно увидеть список работ, процитировавших данную публикацию. Осуществляется эта услуга через *doi*, поскольку, по условиям договора, издатель, кроме обязательной регистрации *doi* статей трех последних лет, в течение 18 месяцев должен разместить у себя на сайте списки литературы этих статей с исходящими ссылками на оригинальные публикации. Таким образом, если статью из Вашего журнала цитируют в другом издании, у которого есть *doi*, высока вероятность того, что читатель с сайта того журнала перейдет на Ваш сайт, а это способствует привлечению новых читателей, авторов и рецензентов. Для участия в *CytedBy* издатель должен подать эти списки литературы (опубликованные на своем сайте) в *CrossRef*. Так информация (по запросу) поступит журналам, которые подписали договор с *CrossRef* и которых процитировали у Вас, следовательно, они смогут добавить данную информацию себе на сайт. Для этого на сайте журнала необходимо предусмотреть возможность размещения и списков литературы, и информации о том, кем данная статья процитирована.

CrossCheck. Система поиска плагиата на сервисе *iThenticate*. После подписания договора и оплаты взноса необходимо загрузить полные тексты вышедших статей в систему. Далее (на этапе подачи статьи) ее текст отсы-

ляется для проверки и, по полученному результату (таблице с процентами идентичности/схожести текстов и источников), редколлегия принимает решение отклонять статью или отсылать на доработку или рецензирование. *Crosscheck* так же является платным сервисом. Оплата формируется из двух типов платежей годового взноса: 1) 20 % от годового взноса в *CrossRef* и 2) постатейная оплата, которая существенно ниже по сравнению с использованием системы *iThenticate* напрямую, существует также система скидок.

CrossMark. Проще всего описать, для чего создан данный сервис, используя гипотетическую ситуацию. После открытия антибиотиков, т. е. с 1940-х годов, их успешно применяли в терапии инфекционных заболеваний. Но у микроорганизмов быстро возникает устойчивость. Поэтому антибиотики постоянно совершенствуются, проходят клинические испытания, рекомендуются для использования в клинической практике и, соответственно, в медицинских журналах публикуются статьи. Врач может прочитать статью, применить устаревшую схему лечения и не получить положительного результата. Однако схема описана в солидном издании, статья найдена в серьезной базе, но не учтено, что появились обновления и уточнения к данной публикации, например, установлено, что бактерии уже выработали механизм резистентности к данной модификации антибиотика. Естественно, ситуация с выбором схемы лечения через Интернет маловероятна в реальной жизни, но для любых исследований важно использовать наиболее точные и "свежие" данные, чтобы не изобретать велосипед. Потому, если на сайте у статьи Вы видите значок *CrossMark*, не поленитесь зайти и проверить, не вышли ли обновления к данной публикации, чтобы цитировать лишь актуальную литературу.

Подводя итоги:

1. Для корректного цитирования журнала важно наличие согласованного с *ISSN* названия журнала на английском языке, которое необходимо разместить на сайте издания.

2. Журналы должны стремиться к включению в базы данных *Scopus*, *elibrary*, *Web of Science* и др. Возможно, индексирование определенными базами должно быть обязательным для изданий, в которых публикуются результаты диссертационных работ.

3. Украинские издания, особенно входящие в " список ВАК", обязаны иметь веб-сайт, на котором должны быть хотя бы содержания архивных номеров всех выпусков на английском языке. Понимая сложность реализации этой задачи для небольших редакций журналов, необходимо создать группу технической поддержки редакций, в работу которой можно привлекать институты академии, специализирующиеся на реализации подобных проектов. Возможно, координировать это сможет Научно-издательский совет Национальной академии наук Украины.

4. Необходимо пропагандировать использование *doi*, проводить ознакомительные семинары, организовать техническую поддержку редакций.



Поскольку НАН Украины является соучредителем академических изданий, очевидно, префикс *doi* академических изданий должен быть зарегистрирован и оплачен академией, а суффиксы – каждым журналом самостоятельно.

5. Пошаговая инструкция для получения индекса *doi*. Издатель должен:
подписать договор с *CrossRef*,
оплатить годовой взнос,
получить префикс,
создать "целевые" страницы,
сгенерировать суффиксы и
подать *doi* статей трех последних лет в *CrossRef*,
в течение 18 месяцев разместить на сайте журнала пристатейные списки литературы с исходящими гиперссылками,
вовремя оплачивать счета.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. <http://www.crossref.org/>
2. <http://www.doi.org/>
3. CrossRef Annual Member Meeting, November 13-14, London, UK