

eLIBRARY DOCUMENT NUMBER (EDN)

eLIBRARY Document Number (EDN) - это уникальный код, который присваивается всем документам на платформе eLIBRARY.RU. Его использование и распространение поможет более эффективно решать целый ряд проблем, возникающих при работе с научной информацией.



ПОСТОЯННЫЙ ИДЕНТИФИКАТОР

Создание постоянного идентификатора для научных публикаций было одной из основных задач при разработке системы DOI. В основу системы была заложена ответственность и заинтересованность издателей в присвоении кодов DOI для своих изданий и поддержании ссылок на полные тексты в актуальном состоянии. Сами агентства по регистрации DOI собирают и хранят только метаданные публикаций и ссылку на полный текст на внешних ресурсах, но не хранят полнотекстовых файлов. Такое распределение ответственности делает систему достаточно гибкой, но приводит к проблеме устаревших ссылок. Постоянные изменения структуры сайтов издателей, реорганизация учредителей или издателей, смена кадрового состава, отказ техники и множество других причин приводят к тому, что ссылки на документы, загруженные ранее в систему DOI, перестают работать. Особенно это касается архивных документов, которые иногда просто исчезают с сайтов издателей. К сожалению, далеко не все издатели отслеживают эти изменения и своевременно актуализируют данные в системе DOI.

EDN, в отличие от DOI, использует гибридную модель ответственности за качество и актуальность данных. Издатель, как и в системе DOI, может полностью контролировать процесс присвоения идентификаторов для своих публикаций, вносить изменения в метаданные и ссылки на полные тексты. Однако, поскольку на платформе eLIBRARY.RU в подавляющем большинстве случаев размещаются не только метаданные, но и полные тексты публикаций, пользователь, запросивший документ по коду EDN, в любом случае получит возможность доступа к полному тексту, даже если ссылка на сайт издательства не будет работать. Полнотекстовые файлы на платформе eLIBRARY.RU размещаются согласно лицензионным соглашениям без ограничения срока, поэтому гарантируется сохранность и доступность архивов в долгосрочной перспективе.

ПОЛНОТА ОХВАТА ИДЕНТИФИКАЦИИ

На данный момент DOI присваиваются не всем типам научных документов. Преимущественно это статьи в журналах, монографии, сборники статей и труды конференций. Некоторые научные документы (например, научные отчеты, диссертации, патенты, препринты и т.д.), как правило, не получают кодов DOI. Для них существуют собственные системы регистрации с присвоением номеров или идентификационных кодов. Однако перейти на полный текст такого документа, встретив его номер в ссылке, не просто и не всегда возможно, не говоря уже о том, что такого номера в ссылке может и не быть.

Кроме того, если говорить о российских издателях, то дополнительными препятствиями для распространения DOI являются платность получения кодов и необходимость оплаты услуг регистрирующих агентств в валюте или работы через посредников, что, естественно, приводит к удорожанию цены. Так, например, лишь немногие из российских издателей готовы потратить деньги на присвоение кодов DOI всем архивным выпускам своих журналов, хотя статьи в этих выпусках могут еще активно цитироваться в текущих публикациях.

Наконец, часть российских организаций сейчас просто не имеет возможности работать с CrossRef, ни напрямую, ни через посредников, из-за санкций, наложенных западными странами. При этом существующие на данный момент времени системы грантовой поддержки и отчетности научных организаций требуют наличия DOI для всех журнальных статей.

С точки зрения полноты охвата информационного потока EDN имеет значительные преимущества. Во-первых, коды EDN присваиваются всем типам научных документов. Во-вторых, эти коды автоматически присваиваются всем новым документам на платформе eLIBRARY.RU, даже если этим не занимается издатель. Это означает, что любой документ, существующий на eLIBRARY.RU, имеет код EDN. Ну и наконец, получение кодов EDN бесплатно, также как и регистрация метаданных на eLIBRARY.RU.

СВЯЗЬ МЕЖДУ ССЫЛКАМИ И ЦИТИРУЕМЫМИ ДОКУМЕНТАМИ

Одной из основных целей, которые ставили перед собой разработчики DOI, было упрощение идентификации ссылок в списках цитируемой литературы. Это действительно очень полезная функция, но она предполагает, что каждая ссылка в списке дополняется кодом DOI. Для нахождения соответствующего кода

DOI в CrossRef существует специальный поисковый интерфейс и API. Однако лишь немногие из российских издателей контролируют качество списков литературы и следят за тем, чтобы ссылки сопровождалась кодами DOI. Кроме того, как уже отмечалось выше, далеко не все цитируемые документы вообще имеют коды DOI.

В EDN эта задача решается похожим образом. Для идентификации ссылок разработан API. Отправив текст ссылки в качестве входного параметра запроса к API, можно на выходе получить коды EDN, DOI, ISBN и eLIBRARY ID при условии, что данная публикация имеется на портале eLIBRARY.RU. Доступ к данному API открыт для всех организаций и частных лиц, требуется только регистрация для получения кода доступа apiKey.

АВТОМАТИЧЕСКАЯ ОБРАБОТКА ДАННЫХ

Важным фактором, влияющим на эффективность использования системы идентификации научных документов, является машиночитаемость кодов, то есть насколько данный код приспособлен для автоматической обработки данных: извлечения этого кода из текста ссылки или библиографического описания, проверки его достоверности и т.д. К сожалению, изначально заложенные принципы формирования кодов DOI сильно ограничивают эти возможности.

Во-первых, поскольку структура кода формируется самим издателем, проверить его достоверность можно только после загрузки этих данных в CrossRef. При этом часто встречается ситуация, когда коды уже напечатаны на страницах выпуска журнала, но еще не зарегистрированы в CrossRef. Соответственно, при распространении метаданных возникает проблема: проверить можно только префикс DOI, а верификация остальной его часть практически не возможна.

Во-вторых, коды бывают очень длинные и могут включать различные символы, используемые в качестве разделителей в библиографической записи (дефисы, точки и т.д.). В результате, в процессе верстки или распознавания текста могут возникать проблемы: расщепление кода, перенос его на новую строку, замена дефиса на тире и т.д. При разработке EDN мы постарались минимизировать эти проблемы. Код EDN, во-первых, короткий (всего 6 символов), во-вторых, может содержать только латинские буквы, и, самое главное, надежно верифицируемый. В системе EDN код не формируется самим издателем, а резервируется из пула свободных кодов через специальный интерфейс. Это означает, что в любой момент времени можно проверить достоверность кода путем запроса через API eLIBRARY.RU. Эта возможность открыта для всех и не требует регистрации и apiKey.