

Цифровая идентификация объектов:

технология и не только

Москва - 2016

УДК 347.78

ББК 67.404

Книга публикуется под лицензией Creative Commons CC BY-NC-SA 4.0.

*Книга подготовлена по заказу Фонда содействия развитию интернета «Фонд поддержки интернет» под общей редакцией **М. А. Медриша**. В написании книги участвовали: **Д.М. Белявский, С.С. Дарбинян, И.И. Засурский, К. Р. Казарьян, И.Ю. Лёвова, В.В. Харитонов**. Технический редактор – **И.Б. Пыжова**.*

Аннотация

Книга «Цифровая идентификация объектов: технология и не только» подготовлена коллективом авторов, каждый из которых внес свою лепту в исследование инструментов цифровой идентификации объектов. В книге дан подробный обзор существующих на сегодняшний день технологий и систем идентификации объектов различной природы, изучены особенности их применения, представлена историческая ретроспектива возникновения таких инструментов. Основное внимание при этом уделяется технологии DOA (Digital Object Architecture) и её частному случаю – DOI (Digital Object Identifier), которые являются сегодня наиболее перспективными и при этом до сих пор недостаточно полно исследованными. Перспективность DOA является следствием изначальной ориентации этой технологии на идентификацию цифровых объектов в сети Интернет. Вторым важнейшим свойством DOA является универсальность, под которой понимается отсутствие ориентации на объекты определённого типа. Именно тесная взаимосвязь с интернетом и универсальность делают технологию DOA наиболее перспективной технологией цифровой идентификации объектов.

ISBN 978-5-9906425-4-6

ISBN 978-5-9906425-4-6



9 785990 642546

© Фонд содействия развитию интернета
“Фонд поддержки интернет”, 2016
© Оформление. **АНО Издательский Дом**
«Научное обозрение», 2016

Введение

Как высказывался профессор Лоуренс Лессиг, компьютерный код в цифровой среде часто описывает идеи и принципы, основанные на социальных или юридических конструктах, принципы субъективные или размытые, зависящие от контекста.

Как можно идентифицировать цифровой объект? Цифровые объекты – всего лишь цепочки битов, которые вполне можно сверить алгоритмически с эталонным образцом. Но этого может быть недостаточно для идентификации цифрового объекта или отражения объекта материального мира в сети Интернет. Человек воспринимает не биты, а их отображение в привычной для человека форме с помощью аппаратного и программного обеспечения. И чем больше в этом процессе задействуется сенсорный аппарат человека (от чтения простого текста к восприятию сложных аудиовизуальных произведений, например, компьютерных игр), тем сложнее и субъективнее вопрос, что считать сущностью цифрового объекта и как его идентифицировать. Идентичны ли две версии цифровой книги с разным форматом файла? С различной версткой? В зависимости от контекста, в одном случае нам будет достаточно эквивалентности текста, в других мы будем сравнивать сигнатуры объектов, созданные при помощи криптографических алгоритмов. Таким образом, определение семантики и контекста среды ложится на другие приложения или механизмы.

Цифровые подписи, связанные с публичной инфраструктурой открытых ключей (PKI), с некоторой уверенностью (в зависимости от доверия к центру сертификации) говорят нам о принадлежности и/или неизменности состояния объекта. Однако для определения контекста подписи по отношению к цифровому объекту необходим определенный словарь терминов и свойств: для архивов, для интеллектуальной собственности, для коммерции и юриспруденции и т.д.

Стороны должны доверять и системам регистрации и поиска объектов. Для этого система должна обладать устойчивостью от действий «плохих» игроков и контролировать возможность изменения информации об объектах. Наличие политики конфиденциальности также необходимо для ограничения возвращаемой информации в зависимости от идентичности запросившего, особенно когда нам требуется обеспечить возможность анонимного доступа. Наконец, все объекты должны быть зарегистрированы под уникальными и, что очень важно, постоянными идентификаторами, назначаемыми доверенным центром (регистратором), а каждому идентификатору должен соответствовать набор метаданных – детальных сведений об объекте, зависящих от контекста. При этом набор метаданных сам, по сути, является цифровым объектом, имеющим ясную структуру.

Таким образом, идентификация и аутентификация цифровых объектов требует не только наличия технических алгоритмов, но и участия социальных структур для создания доверенной структуры, общепринятой терминологии и обеспечения совместности и взаимодействия между различными системами.

Раздел 1.

Модели идентификации цифровых объектов в интернете: стандарты и перспективы

Цифровые объекты

Цифровая сущность (или цифровой объект) – ключевая концепция международных рекомендаций МСЭ-Т X.1255 «Структура обнаружения информации по управлению определением идентичности», общепринятая структура данных, состоящая из одного или нескольких элементов, благодаря которой обеспечивается функциональная совместимость (интероперабельность) информационных систем в интернете.

Для управления цифровыми объектами используются три архитектурных компонента. Каждый из компонентов может использоваться самостоятельно, но в комбинации они обеспечивают распределенную и масштабируемую систему управления информацией в интернете. Эти компоненты таковы:

- а) масштабируемая и распределенная система идентификаторов и резолюции цифровых объектов;
- б) репозитории доступа и управления цифровыми объектами;
- в) реестры для поиска и обнаружения объектов.

Цифровые объекты – ключевой элемент, вокруг которого выстроены другие компоненты и сервисы. Цифровые объекты не заменяют существующие форматы и структуры данных, но обеспечивают общепринятые способы представления этих форматов и структур. Это позволяет их однозначно интерпретировать

и перемещать между различными гетерогенными информационными системами в ходе изменений в системах по прошествии времени. Таким образом, компоненты системы позволяют долгосрочное управление структурированной в цифровые объекты информацией, уникально идентифицируя их на постоянной основе, обеспечивая способ получения текущего состояния объектов и способ получения объектов или их метаданных. При этом система резолюции цифровых объектов должна отвечать требованиям производительности, быть распределенной и главное – поддерживать неизменность ссылок на протяжении жизненного цикла системы, вне зависимости от смены местоположения объектов, способов доступа, владения или других атрибутов. Идентификатор (или хэндл) цифрового объекта используется для определения информации о состоянии самого объекта. Администратор объекта может включить в такую информацию данные о расположении объекта, методах аутентификации, публичные ключи и другие данные, позволяющие работать с информацией.

Поскольку цифровой объект по сути – это цепочки битов, которые могут быть уникально идентифицированы, то при этом часть объекта также может быть идентифицируемой. В контексте рекомендаций МСЭ-Т объектами могут быть сети, сервисы, документы, права доступа, транспортная информация, устройства или отдельные чипы, авторские произведения или любая другая информация, представленная в виде структуры данных и связанных с ними метаданных, то есть в виде цифрового объекта. Таким образом, в качестве цифрового объекта можно представить и любой объект реального мира.

Идентификация объектов в интернете

Изначальная архитектура интернета создавалась для связывания сетей пакетной передачи данных и вычислительных ресурсов, объединенных этими сетями. Глобальный механизм адресации, основанный на протоколе IP, предполагал, что каждому устройству будет присвоен свой уникальный адрес, состоящий

из номера сети и номера устройства в этой сети. Таким образом, идентификатор зависел от местоположения, а не от содержания пакета или метаданных. С развитием интернета и появлением системы DNS стало возможным перейти от цифровых идентификаторов ресурсов в пользу текстовых адресов URL. Идентификатор в такой системе все еще зависит от местоположения, но уже появилась возможность указывать на точный адрес объекта, а с помощью расширений протокола HTTP (MIME) и на тип объекта.

Вместе с тем, как следует из той же рекомендации X.1255, для согласованного долговременного управления информацией в интернете необходимо сочетание модели данных, протокола для взаимодействия и системы идентификаторов. Подобная система идентификаторов должна обладать определенными свойствами, такими как уникальность, постоянство, возможность резолюции, интероперабельность. Система идентификации и резолюции идентификаторов должна быть распределенной, надежной, высокопроизводительной, масштабируемой и глобальной. Необходимо обеспечивать постоянство и неизменность ссылок на объекты (по их идентификаторам) в течение длительных периодов времени, вне зависимости от изменений местоположения, методов доступа, владения и других изменяемых атрибутов метаданных. Наконец, технологии должны быть применимы к объектам любого рода: материальным, цифровым, логическим.

Существующие до сих пор системы идентификации в основном относились к конкретной группе объектов или системе контроля над объектами. Для цифровых объектов же необходима система, нейтральная по отношению как к природе объекта, так и прикладной цели идентификации. При этом она должна позволять связывать объекты с интернетом, идентификаторы объектов – с метаданными, а также обеспечить возможность изменять и обновлять метаданные.

Возвращаясь к принципам адресации в интернете, к примеру, URL указывает на местоположение файла, но не идентифицирует объект и не обеспечивает постоянство ссылки, и, хотя

предпринимались попытки исправить эти недостатки стандартами URN/URI, они не содержат полноценной схемы идентификации объектов.

Исходя из многообразия пользовательских и бизнес-практик, можно определить основные требования к технологиям идентификации, учитывающие использование интернета. Для начала стоит определиться, что *идентификация необходима*. Каждая сущность, которую необходимо распознавать в цифровых сетях, должна иметь хотя бы один публичный или совместно используемый идентификатор. Публичный идентификатор обязательно должен быть человекочитаемым, он должен быть доступен через сеть.

Идентификаторы должны обнаруживаться (быть разрешимыми). Разрешимый идентификатор в цифровой сети позволяет системе найти идентифицируемый объект или информацию о нем (т.е. метаданные). При этом необходимо обеспечить поддержку Сети, WWW и совместимость с URI как основным форматом взаимодействия. Синтаксис URI позволяет включать существующие стандарты идентификаторов, сохраняя уникальность и разрешимость.

Идентификаторы должны отвечать на множественные запросы. Идентификатор может указывать на различные местоположения в зависимости от типа метаданных. Множественная резолюция должна быть возможна без дополнительных знаний со стороны пользователя, используя стандартные протоколы. Подобные качества, в том числе, позволяют обеспечить совместимость открытых систем и устаревших или проприетарных систем идентификации. Необходимо учитывать, что для эффективной работы множественной резолюции необходимы стандартизированные словари для описания метаданных объектов.

Для работы с идентификаторами необходимы различные уровни доступа. Регистрация идентификаторов должна проходить под контролем доверенных центров, с обозначенными политиками безопасности, использования и доступа. В системе идентификации должна существовать достаточная поддержка описательных данных и данных о владении объектом

в метаданных, при этом регистрацией метаданных должна заниматься доверенная структура: так как метаданные могут быть определены многочисленными пользователями, возможно возникновение конфликтов, для разрешения которых и потребуется арбитр. При этом существование распределенных механизмов доверия критично важно для функционирования системы и предотвращения дублирования или конфликтов идентификаторов.

Владение метаданными должно быть отделено от объекта. Пользователь, задающий метаданные, может и не быть владельцем или издателем данных, а только посредником, которому, тем не менее, требуется обновлять метаданные из различных источников, порой конфликтующих между собой. Система должна предоставлять гибкие возможности по администрированию сведений об объектах.

Идентификаторы не должны содержать динамические элементы или метаданные. Как правило, «глупые» идентификаторы, то есть идентификаторы, содержание которых не несет смысла само по себе, предпочтительнее для здоровья системы. Но в некоторых случаях ограниченная информация, содержащаяся непосредственно в идентификаторе, может нести пользу или даже необходима. Например, информация о типе самого идентификатора (dx.doi.org; ISBN). Информацию о создателе и дате публикации лучше не включать в идентификатор, чтобы избежать ошибок при чтении человеком (который может спутать владельца и создателя, дату создания с датой публикации). Тем не менее, многие стандарты идентификации объектов включают подобные сведения. Неизменная информация об объекте не должна включаться в идентификатор, также как и любые другие метаданные, она может быть интерпретирована по-разному, в зависимости от контекста, и, поскольку со временем она может стать некорректной, подобные данные должны включаться в метаданные, которые будут доступны в резолюции идентификатора. Динамическая информация никогда не должна включаться в идентификатор, данные принципы также относятся и к цифровым отпечаткам объектов.

Метаданные и требования к описанию объектов

Существует выражение, что метаданные – это источник жизненной силы для электронной коммерции. Электронная торговля чрезвычайно сильно зависит от процедур идентификации продавцов, покупателей и товаров, а также способов описания товаров. Электронная коммерция нуждается в идентификаторах, которые бы связывали людей с товарами и услугами. Требования к ним значительно усложняются, если речь идет о коммерческих процессах, связанных с интеллектуальной собственностью, где отношения между участниками сделок могут быть весьма запутанными: существуют различные типы авторских и смежных прав, формы лицензирования, не говоря уже о средствах воспроизведения или географических ограничениях. Любая сущность должна быть уникально идентифицирована согласно принятому стандарту и описана метаданными.

Эффективность многих сегментов коммерции зависит от того, насколько развиты в них системы идентификации: достаточно взглянуть на штрих-коды UPC или на идентификаторы ISBN, которые в течение долгого времени являлись важнейшей частью инфраструктуры дистрибуции в соответствующих отраслях. В то же время в тех сегментах, где уникальные идентификаторы не существуют или слабо развиты, издержки на управление данными и управление вообще куда выше. Ярким примером можно считать ситуацию с аудиоконтентом, где стандарт ISRC так и не получил широкого распространения, или с другими типами контента и интеллектуальной собственности.

В реальном мире описание объекта часто включает разнообразные данные и метаданные, которые должны быть описаны в структурированной манере, так же, как и тип объекта, а также все возможные действия над объектом. Для использования метаданных в мультимедийной, мультязыковой среде необходимы уникальные идентификаторы на всех уровнях, включая словари типов объектов и их атрибутов. В идеальной с точки зрения прикладного использования ситуации единственное свободно

определяемо поле в метаданных должно быть именем или названием, но в некоторых случаях и его значения могут быть ограничены, например, зарегистрированными торговыми знаками или законодательством.

С другой стороны, некоторые проблемы, относящиеся к «офлайновым» идентификаторам, не так важны для цифровых. Так в цифровых идентификаторах нежелательно включение метаданных в формат идентификатора из-за ненадежности подобного подхода. В то же время в цифровой среде одна и та же сущность может иметь несколько идентификаторов, например, в зависимости от формата или языка, или другой практической необходимости гранулярно идентифицировать тот или иной объект.

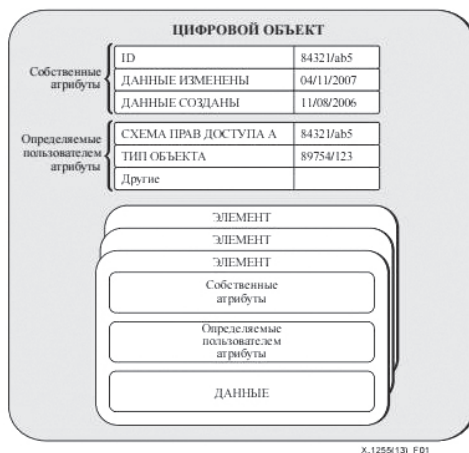


Рисунок 1. Структура цифрового объекта в соответствии с рекомендациями X.1255

Метаданные

Международный стандарт ISO/IEC 11179 (Реестры метаданных) описывает семантику данных, представление данных и упорядочивание описаний этих данных. Именно благодаря метаданным и возможно точное понимание семантики данных для использования их в дальнейшем. Благодаря метаданным становится возможным стандартизированное описание данных, общее понимание данных внутри организационных структур и

между различными организациями; метаданные поддерживают переносимость и интероперабельность данных и приложений.

Согласно стандартам ISO, метаданные – это данные, которые описывают и определяют другие данные. Набор условий или целей, при которых данные используются как метаданные – это контекст. *То есть метаданные – это данные о данных в некотором контексте.* Понимание данных фундаментально для их использования, обмена и стандартизации. Концептуальная семантика описывает типы объектов, для которых собираются данные, и конкретные характеристики описываемых объектов. Символьная семантика описывает множество значений, которые могут принимать те или иные атрибуты объекта. Одним из условий корректного использования данных является общее понимание значения и описания данных как пользователями, так и владельцем данных.

Стандартное определение метаданных, однако, не говорит о том, откуда они берутся или как их выделять. В документе ISO/IEC 11179 приводится следующий пример:

Допустим, в ячейке таблицы содержится следующая строка: 2013-12-10-14:30:00. Допустим также, что эта строка означает дату и время взлета детского волшебного дракона – 10 декабря 2013 года в 14.30. Этот факт – интерпретация, значение, скрывающееся за данными, результат применения инфологического уравнения к исходным данным. Вместе с тем, в другое время и при других условиях возможна другая интерпретация этих данных. Смысл факта – это информация, которую несут данные, и человек интерпретирует данные, чтобы его постичь. Но смыслы и информация – это идеи, которые люди хранят в голове. Строка символов, значение которой мы назвали фактом – это материализация этого смысла. Кроме того, этот факт является описанием ячейки данных, а описание передает смысл некоего объекта. Наше знание об объекте может быть неполным, но добавление контекста может помочь составить полную картину. Так дата и время, погода в месте события, методика и точность измерения могут быть релевантными для интерпретации. Каждая частица информации может быть добавлена к факту по мере

получения. Например, «Датой и временем взлета детского волшебного дракона в прохладный облачный день на Черном озере в Школе Чародейства и Волшебства Хогвартс в Шотландии, согласно магическим наблюдениям с Башни Гриффиндора, было 10 декабря в 14.30 по Гринвичу».

Данное высказывание является описанием и смыслом представленной ячейки данных. Это высказывание также является предложением в документе, то есть представлено в виде данных само по себе. Таким образом, это высказывание – метаданные. Если мы хотим представить высказывание в виде строки базы данных, стандарт ISO/IEC 11179-3 описывает метамодели для организации метаданных.

Класс объекта – взлет детского волшебного дракона

Свойство – дата и время

Домен значений – описание даты и времени арабскими цифрами

Формат – год-месяц-день:час-минута-секунда

Точность – секунда

Тип данных – дата-время

Некоторые метаданные, такие как погода, имя и местоположение школы, не являются стандартными, так что для их представления необходимо определиться со словарем атрибутов.

Системы идентификации объектов

В мире существует множество различных систем идентификации объектов, которые можно условно разделить на несколько типов:

- Нишевые – отраслевые системы идентификации, или системы, созданные для использования в конкретном проекте, например:

- Книги – ISBN(ISO2108)
- Авторы – ISNI(ISO27729), DAI, ORCID
- Медицинские публикации – PMID
- Информация любого типа – ARK

- Стандартизированные – принятые на уровне международных организаций, таких как ISO, МСЭ, IETF, GSMA, 3GPP и других:

- Объекты в интернете – URN (RFC2141), URI(RFC3986)
- Виртуальные объекты – OID (ITU-TX.690), UUID (RFC4122)
 - Цифровые объекты – DOI (ISO 26324)
- Национальные – используемые и продвигаемые отдельными государствами или наднациональными объединениями:
 - Товары и продукты в ЕС – EPC, EAN
 - Интернет вещей в Китае, Японии – CID, ECODE, ucode
- Глобальные – принятые к использованию во всем мире:
 - Книги – ISBN (ISO2108)
 - Автомобили – VIN(ISO3779)
 - Мобильные телефоны – IMEI (3GPP TS 22.016)
- Устаревшие или не получившие распространения:
 - Ресурсы в интернете – PURL, XRI

Данная классификация является во многом условной, и практически любую систему можно отнести сразу к нескольким категориям, однако стоит отметить влияние стандартизации на распространенность той или иной системы идентификации в мире.

PMID

PMID – это акроним идентификатора PubMed (PubMed – бесплатная поисковая система, обеспечивающая доступ к базе данных MEDLINE, где содержатся цитаты и выдержки из статей биомедицинских исследований) или, точнее, уникальный номер, присваиваемый каждому ресурсу PubMed, такому как цитата, исследование в области биологических наук или биомедицинская научная статья в рецензируемых журналах. Данный акроним используется аналогично Международному стандартному книжному номеру (ISBN) для книг. В 2005 году насчитывалось примерно 15-16 млн. номеров, начиная с PMID равного 1. В настоящее время около 1 млн. новых номеров добавляются каждый год. С технической точки зрения Уникальный идентификатор [UID] является специальным полем-тегом, используемым в поисковых запросах по сайту. Поиск по статьям с использованием PMID и с соответствующими аргументами (с, без [uid] или тега области) возвращает синопсис подходящих по критериям работ. Допускается одновременный поиск нескольких

публикаций по их PMID (с уточняющим тегом [pmid] или без него), введенным в поисковую строку PubMed через пробел. При одновременном поиске названий публикаций и других терминов использование стандартных поисковых тегов PubMed обязательно: lipman[au] 16381840[pmid]. Идентификаторы PMID, присвоенные публикациям в PubMed, впоследствии никогда не меняются и не используются повторно.

ARK (АРК)

Схема АРК (Archival Resource Key) предназначена для идентификации информационных объектов различного типа на постоянной основе. Идентификатор ARK имеет следующую структуру: [http://NMAN/]ark:/NAAN/Name[Qualifier]. В данной схеме NAAN – это префикс регистрационного агентства, NMAN – это префикс текущего места хранения объекта (или метаданных о нем), а Qualifier – опциональный компонент, позволяющий указывать на подкомпоненты или различные версии объекта. Таким образом, идентификатор модифицирует стандартную схему URL. ARK – это адреса, предназначенные для поддержки долгосрочного доступа к информационным объектам, доступ к которым требуется существенно больше, чем ресурс, предоставляющий доступ к нему. ARK идентифицируют объекты любого типа:

1. Цифровые объекты – документы, базы данных, изображения, программное обеспечение, веб-сайты и т.д.
2. Физические объекты (книги).
3. Живые объекты и группы: людей, животных, компаний, оркестров и т.д.
4. Геолокация, заболевания, словари, и т.д.

ARK удобен по целому ряду причин:

1. Доступность – нет сборов для назначения или использования ARK.
2. Самодостаточность – вы можете разместить ARK на собственном веб-сервере.
3. Переносимость – вы можете переместить ARK на другие серверы без потери своих основных идентичностей.
4. Возможность ARK обрабатывать смешанный регистр, позволяющий использовать более короткие идентификаторы.

Вместе с тем, система ARK не предназначена для глобального использования или масштабируемости: так, хранение идентификаторов по сути происходит в локальных базах данных, а для доступа к ним используются скрипты на Perl и т.п. За последние четырнадцать лет больше 395 организаций распределены в пятнадцати странах зарегистрированных ARK. Зарегистрированные организации владеют библиотеками, архивами, музеями (Smithsonian), издателями (PeerJ), государственными учреждениями (EPA), научными учреждениями (Принстон). Некоторыми из основных пользователей являются:

- Цифровая библиотека Калифорнии
- Internet Archive
- Национальная библиотека Франции
- Калифорнийский университет в Беркли
- Университет Северного Техаса
- Чикагский университет
- University College Dublin
- Британская библиотека
- Google

В 2013 году схему идентификации ARK пытались сделать стандартом IETF, однако он не был принят.

IMEI

Международный идентификатор мобильного оборудования – число (обычно 15-разрядное в десятичном представлении), уникальное для каждого использующего его аппарата. Применяется в сотовых телефонах сетей GSM, WCDMA и IDEN, а также в некоторых спутниковых телефонах. IMEI присваивается уполномоченными организациями, в частности, ВАВТ. IMEI (14 десятичных цифр плюс контрольная цифра) содержит информацию о происхождении, модели и серийном номере устройства. Первые 8 цифр составляют модель и место происхождения устройства и известны как ТАС (Type Approval Code). Остальная часть – определяемый производителем серийный номер аппарата с высчитанной по алгоритму Луна контрольной цифрой в конце. До 2003 года эта цифра обязательно должна была равняться 0. Позже это правило было отменено, например, нынешний iPhone 5 имеет ТАС

с цифрами 01-332700, а Samsung Galaxy S2 использует 35-853704. Секция кода, помеченная как буква С, является уникальным серийным номером телефона и проставляется исключительно производителем. Последняя цифра является контрольной и используется для проверки всей строки. IMEI вшит в устройство и не имеет никакого отношения к SIM-карте. Если ваш телефон украли, настройки сброшены, а SIM-карта заменена, номер IMEI не изменится без постороннего вмешательства. Как правило, IMEI указывается в четырех местах: в самом аппарате (в большинстве случаев его можно вывести на экран набором *#06# на клавиатуре), под аккумуляторной батареей, на упаковке и в гарантийном талоне. IMEI играет роль серийного номера аппарата при авторизации в сети, передается в эфир при авторизации в сети. Также IMEI используется для слежения за аппаратами и блокирования краденых телефонов на уровне оператора сотовой связи, что не позволяет в дальнейшем использовать такой аппарат в сети этого оператора, однако не мешает его использованию в других сетях.

UPC/UPC/EAN

UPC: американский стандарт штрих-кода, предназначенный для отслеживания товаров в магазинах. UPC был разработан в 1973 году Джорджем Джозефом Лорером, работавшим инженером в корпорации IBM. Стандартизацией и регистрацией кодов UPC занимались организации UCC (Uniform Code Council, Inc.) в США и ECCC (Electronic Commerce Council of Canada) в Канаде. В 2005 году эти организации объединились с европейской ассоциацией EAN и образовали глобальную организацию по стандартизации GS1. «Национальная организация GS1 в России» является представителем этой организации в нашей стране.

Разновидности кода:

- UPC-A (полный) – кодируется 12 цифр.
- UPC-E (сокращенный) – кодируется 8 цифр.

Код UPC содержит только числа и никаких букв или других символов. Код UPC – простой и практически симметричный линейный штрих-код. Простота, симметричность и высокая помехозащищенность обусловлены недостаточно развитой техникой времен создания этих кодов.

Код состоит из 2 групп цифр, по 6 цифр в каждой группе – левой и правой. Группы цифр окаймляются так называемыми защитными (или ограждающими) штрих-шаблонами (Guard Patterns). Эти шаблоны содержат штрихи единичной ширины, которые служат для синхронизации сканера штрих-кода. Наличие именно трех таких полей обусловлено, в первую очередь, возможным нанесением штрих-кода на закругленную поверхность. И если сейчас это не является особой проблемой, то во времена создания этого кода сканеру требовалось знать ширину единичного штриха в начале, середине и конце кода. Левые и правые защитные шаблоны состоят из 3 штрихов единичной ширины – двух темных и одного светлого между ними. Средний защитный шаблон состоит из 5 штрихов – трех светлых и двух темных. Все остальное – цифры.

Каждая цифра левой или правой группы кодируется с помощью четырех штрихов: двух светлых и двух темных. Каждый штрих может иметь относительную ширину в одну, две, три или четыре единицы. Общая ширина штрихов для одной цифры всегда составляет семь единиц. Битовая комбинация для каждой цифры разработана таким образом, чтобы цифры, насколько это возможно, отличались друг от друга. Максимальная длина темного или светлого участка не может превышать четырех единиц. Общая ширина всего кода всегда равна 95 единицам. В любом коде 29 светлых и 30 темных штрихов. Все эти технические решения важны для надежности и простоты сканирования этого кода.

Первая цифра кода – это так называемый префикс – имеет некоторое логическое значение, но не столь важна с технической точки зрения. Последняя цифра – контрольное число, служит для выявления возможной ошибки при чтении кода сканером или ручного ввода цифр кода с клавиатуры.

Направление чтения комбинации штрихов не имеет значения – код разрабатывался так, чтобы он одинаково просто считывался как в прямом, так и обратном направлении (если товар перевернут). Также не имеет значения то, какое исполнение имеет весь штрих-код – фотографически позитивное или негативное:

штрих-код, нанесенный светлыми полосками по темному фону, читается так же, как и темными полосками по светлому фону. Цвета штрихов и фона не обязательно должны быть белыми и черными, возможны и другие цветовые комбинации.

В коде UPC-A (GTIN-12) контрольное число (цифра) рассчитывается следующим образом:

1. Суммируются все цифры на нечетных позициях (первая, третья, пятая, и т. д.), и результат умножается на три.
2. Суммируются все цифры на четных позициях (вторая, четвертая, шестая, и т. д.).
3. Числа, полученные на предыдущих двух шагах, складываются, и из полученного результата оставляют только последнюю цифру.
4. Эту цифру вычитают из 10.
5. Конечный результат этих вычислений и есть контрольная цифра (десятке соответствует цифра 0).

Например, контрольное число для штрих-кода UPC-A «03600029145X», где «X» – это искомая контрольная цифра, рассчитывается путем сложения всех нечетных цифр ($0+6+0+2+1+5 = 14$), умножается на три ($14 \times 3 = 42$), результат суммируется со всеми четными цифрами ($42+3+0+0+9+4 = 58$), отбрасывается все, кроме последней цифры ($58 \bmod 10 = 8$), вычитается из 10 ($10 - 8 = 2$), и еще раз, если это необходимо, отбрасывается все, кроме последней цифры ($2 \bmod 10 = 2$). Искомое контрольное число – цифра 2.

При считывании кода правильность считывания проверяется похожим способом, но несколько проще:

1. Суммируются все четные цифры и умножаются на 3.
2. Суммируются все нечетные цифры включая контрольную цифру.
3. Эти суммы складываются, и оставляют последнюю цифру от результата.

Технически цифры обрабатываются последовательно, за один проход, с умножением каждой цифры на 1 или 3 в зависимости от четности позиции, добавлением к сумме и взятием остатка по модулю 10 от текущей суммы. Иными словами, десятки сразу

отбрасываются, что сильно упрощает механизм вычисления. Если результат равен нулю, то принимается решение, что код считан правильно, если любая другая цифра, то код однозначно считан неверно.

EAN: европейский стандарт штрих-кода, предназначенный для кодирования идентификатора товара и производителя. Является надмножеством американского стандарта UPC. Стандартизацией и регистрацией кодов EAN занималась европейская ассоциация EAN, в продолжение развития стандартов, разработанных организациями UCC (Uniform Code Council, Inc.) в США и ECCS (Electronic Commerce Council of Canada) в Канаде. В 2005 году эти организации объединились и образовали глобальную организацию по стандартизации «GS1».

Этот стандарт является универсальным для всех стран, а американская и канадская национальные организации настойчиво рекомендовали всем торговым компаниям своих стран заменить устаревшее оборудование и программы UPC, не способные считывать EAN, на современные.

Разновидности кода:

- EAN-8 (сокращенный) – кодируется 8 цифр.
- EAN-13 (полный) – кодируется 13 цифр.
- EAN-128 – кодируется любое количество букв и цифр, объединенных в регламентированные группы.

Коды EAN-8 и EAN-13 содержат только цифры и никаких букв или других символов. Например, 2400000032632. Кодом EAN-128 кодируется любое количество букв и цифр по алфавиту Code-128. Например, (00)353912345678(01)053987(15)051230, где (15) группа обозначает срок годности 30 декабря 2005.

Последовательности штрихов упорядочены по возрастанию ширины штрихов. Цифры со знаком звездочки указывают на L-G схемы кодирования цифр, то есть звездочки нужны для определения самой первой (13-й) цифры. Геометрически ширина черной полоски может быть на единицу больше (по сравнению с шириной пробела) из-за того, что штрих-код иногда печатается на струйном принтере, и краска пропитывает бумагу либо размывается жидкостью.

ЕРС: код ЕРС, или электронный код продукции, представляет собой уникальный, сериализованный идентификатор для объектов любого типа. Формат кода ЕРС позволяет закодировать известные идентификаторы, например:

- GTIN (Global Trade Item Number) – глобальный номер предмета торговли;
- GRAI (Global Returnable Asset Identifier) – глобальный идентификатор возвратных активов;
- SSCC (Serialized Shipping Container Code) – серийный код транспортной упаковки;
- GIAI (Global Individual Asset Identifier) – глобальный идентификатор индивидуального актива;
- CAGE/DoDAAC (US Department of Defense Internal Numbering System) – идентификатор системы внутренней нумерации Министерства обороны США.

Знакомый всем штриховой код EAN/UPC, который используется для кодирования глобального номера предмета торговли, обеспечивает идентификацию класса объектов. ЕРС выходит за эти рамки и позволяет с помощью номера GTIN, включенного в метку ЕРС, уникально и индивидуально идентифицировать каждый объект, принадлежащий классу (таким образом, код ЕРС – это сериализованный глобальный номер предмета торговли, sGTIN). Две одинаковые упаковки одного и того же продукта будут иметь одинаковый номер GTIN, но разные коды ЕРС (сериализованные номера sGTIN). Помимо идентификации потребительских товаров на уровне единичной упаковки, ящика или паллеты, с помощью кода ЕРС можно кодировать такие идентификаторы, как GRAI, GIAI и использовать его для идентификации возвратных контейнеров, зарегистрированного багажа; электроники на предмет предоставления гарантийных и сервисных услуг; для ведения записей эксплуатации основных средств; идентификации контейнеров любого уровня в местной и международной логистике; индивидуального прослеживания емкостей с медицинскими препаратами или любых других материальных объектов.

ЕРСglobal обеспечивает возможность создания индустрии, основанной на стандартах, в рамках которой метки, нанесенные

на объекты, могут перемещаться из одной организации в другую, вплоть до конечного пункта назначения, при этом идентификация и аутентичность объектов будут одинаково понятны везде. Стандарт EPCIS дополняет стандарты EPCglobal и обеспечивает детальное представление о месте нахождения и статусе объекта в процессе его перемещения через границы организаций, а также предоставляет возможность коллективного доступа к этой информации со стороны партнеров и организаций, независимо от применяемой технологии. Следует отметить, что стандарт EPCIS не зависит от носителя данных и может использоваться для обмена данными, полученными при считывании меток RFID, штриховых кодов и других носителей данных.

Штрих-коды и RFID. Идентификаторы для технологий автоматической идентификации (штрих-коды, RFID) широко используются в Европе и находятся на очень развитом уровне. К концу 2000-х основные европейские компании, связанные производством, транспортировкой и розничной торговлей, уже запустили или проходили пробы технологии RFID. Европейцы также запустили несколько программ для дальнейшего развития, применения и расширения технологии RFID (CERP, CASAGRAS12, GRIFS). Недавние исследования показали тенденцию к интеграции RFID идентификаторов с другими типами автоматической идентификации (штрих-коды, QR-коды). Особое распространение такая интеграция имеет место в сервисах IoT (Internet of Things), которые являются вычислительной сетью физических объектов (вещей), оснащенных встроенными технологиями для взаимодействия друг с другом или с внешней средой. При этом наблюдается тенденция к отказу от проприетарных схем идентификации в RFID и использованию открытых стандартов.

PURL

Постоянный единообразный локатор (определитель местонахождения) ресурса. Формат PURL основан на URL – стандартизированном способе записи адреса ресурса в сети Интернет. PURL – это постоянный локатор ресурса, который (в отличие от обычного URL) не может измениться. В базовом виде PURL почти не отличается от обычных URL:

<http://purl.oclc.org/OCLC/PURL/FAQ>

/|\

протокол сервер с БД PURL имя

В PURL нельзя использовать символы ‘#’ и ‘~’. Схема обращения к ресурсу выглядит так:

```
+-----+ PURL +-----+
| | ----->> | |
| | | PURL |
| K | URL | СЕРВЕР |
| Л | <<----- | |
| И | +-----+
| Е | URL +-----+
| Н | ----->> | |
| Т | |СЕРВЕР |
| | РЕСУРС |РЕСУРСА |
| | <<----- | |
+-----+ +-----+
```

Такая схема является не идеальной, и лишнее звено в этой цепи создает дополнительный сетевой трафик и нагружает сеть. Но скорость обработки PURL все равно очень высока – например, еще в 1996 году однопроцессорный файловый сервер «Sun4», обслуживавший базу данных PURL из 500 000 записей, мог перенаправлять 50 запросов в секунду. Разработанный в 1995 году, PURL продвигался как решение, способное удовлетворить потребности в неизменном указателе на интернет-ресурс до выхода полной спецификации URN, и полностью совместимое с ней. С выходом спецификации, а впоследствии и системы DOI, на которую PURL оказал достаточно сильное влияние, использование PURL практически прекратилось. Система Handle оказалась гораздо более приспособленной для масштабируемости и не зависела от работы DNS.

XRI

Расширяемый идентификатор ресурса, разработанный организацией OASIS, прежде всего, как будущая замена URL в

интернете. XRI планировался как новая, совместимая с IRI и URI схема (протокол) для создания абстрактных идентификаторов ресурсов. Такие идентификаторы не зависят от платформы, домена, путей и программ – они полностью абстрактны и поэтому могут быть едины для всех доменов и каталогов. Идентификаторы XRI имеют два слоя: I-Numbers – постоянные сетевые адреса (похожие на IP-адреса). Такие адреса будут регистрироваться на какой-либо ресурс (человека, организацию, приложение, файл, цифровой объект и т. д.) и никогда больше не перерегистрироваться (в отличие от IP-адресов и доменов DNS). Это означает, что идентификатор I-Number всегда можно будет использовать как адрес для какого-либо ресурса (по крайней мере, пока этот ресурс доступен в сети). I-Names – удобные для человека названия (похожие на домены системы DNS). Как имена доменов DNS разрешаются DNS-серверами в IP-адреса, так и I-Names будут разрешаться в I-Numbers. Но, в отличие от I-Numbers, идентификаторы I-Names смогут быть переопределены их владельцами для идентификации других ресурсов. Например, если I-Name ассоциирован с названием компании, а компания решила изменить свое название, то она может передать свой старый идентификатор I-Name другой компании. При этом обе компании останутся со своими старыми идентификаторами I-Numbers.

I NAMES	I NUMBERS
=Mary.Jone @Jones.and.Company +phone.number +phone.number/(+area.code) =Mary.Jones/(+phone.number) @Jones.and.Company/(+phone.number) Jones.and.Company/((+phone.number)/ (+area.code))	!!1002!A7C5 !!1002!A7C5/!D90E.88

Организация W3C отвергла попытки стандартизации XRI, и работа над второй версией системы была свернута. Вместе с тем, многие наработки XRI легли в основу протокола OpenID и стандарта OAuth.

URI

URI (Uniform Resource Identifier) предоставляет возможности для идентификации ресурса во Всемирной Паутине (WWW). Каждый URI начинается с имени схемы, которая, в свою очередь, специфицирует идентификаторы. Например, широко распространенное “http” – это лишь один пример такой схемы (еще существуют mailto, ftp, telnet, file и т.д.). Спецификация URI определяет реализацию доступа к файловому серверу, обычно через http-протокол, хотя может использоваться и другой; и синтаксис для ссылок, с помощью которых, например, ISBN идентификатор можно было бы определить как URI.

URI часто путают с терминами URL и URN. В конце концов, за URI закрепили в RFC 3986 точное значение: либо указатель на местоположение, либо имя, либо оба сразу. С этой точки зрения, термин URL является подклассом URI, который, в дополнение к идентификации ресурса, обеспечивает способ его нахождения; термин URN, в свою очередь, исторически использовался для обозначения имен URI, которые должны оставаться уникальными в глобальной сети и не пропадать, даже если ресурс прекращает существовать или становится недоступным. RFC 3986 объявляет оба термина устаревшими. Но, несмотря на техническую унификацию понятия URI, все равно существует несколько причин для путаницы:

- Цитированные документы, которые ссылаются на позиции, которые существовали ранее, но теперь перенесены.
- Использование общего понятия URI может скрыть полезные различия, которые необходимы для некоторых пользователей, например, библиотекарей, которым может быть полезно различать имя и местоположение, особенно если один и тот же документ располагается в нескольких местах сразу.
- Отношения между несетевыми идентификаторами (ISBN, RFID и т.д.) и URI могут привести к следующему:
 - 1) путанице между идентификатором, представлением и механизмом доступа;
 - 2) недооценке использования идентификатора вне сети;
 - 3) использованию нецифровых символов.

Таким образом, ранее URI имело два подкласса: URN (для идентификации имен) и URL (для идентификации местоположения). После RFC 3986 все сетевые схемы – это схемы URI, и они могут определять как механизм доступа (например, “http:”), так и пространство имен (например, “urn:”).

Некоторые IoT-приложения используют URI для идентификации IoT-объектов и ресурсов. Но некоторые так же определяют типы идентификации в форме UUID. URI поддерживает приложения, которые интегрируют в себе множество идентификационных решений (например, EPC и IPv6). Такая интеграция считается важной для некоторых приложений, которые выходят за рамки нескольких разнородных IoT-систем, использующих различные идентификационные технологии.

URN

Uniformed Resource Name – это спецификация для определения имен (идентификаторов) ресурсов для использования в интернете. RFC 2141 (более старый, чем RFC 3986, объявивший URN устаревшим понятием) определял формальный регистрационный процесс в качестве пространства имен URN и сопровождающую спецификацию для осуществления функциональных требований для пространства имен urn. Существующие идентификаторы могут быть специфицированы как URN, например, urn:isbn: 9789521061547.

Архитектура URN подразумевает сервис (RDS), который мог бы найти подходящий сервис для данной URN схемы. Но глобального RDS не существует: браузеры не могут работать с URN-строками без дополнительных плагинов. Таким образом, не существует никакой гарантии кроссплатформенности: для каждого URN может понадобиться свой плагин, которого для конкретного браузера может не быть. Поэтому большинство современных реализаций URN – это протокол http, который содержит URL релевантного сервиса (например, для примера с ISBN выше, таким сервисом будет Финский национальный URN сервис <http://urn.fi>, и действующая форма URN будет: <http://urn.fi/URN:ISBN:978-952-10-6154-7>). Регистрация URN на сегодняшний момент требует дополнительного

администрирования для определения пространства имен URN (например, строку urn:doi:10.1000/1) и для перенаправления к релевантному сервису.

CID

Система CID (Communication Identifier, идентификатор коммуникации) – это публичная пользовательская система присвоения ID (идентификационных номеров, имен) в контексте Internet of Things. Это система распределения, управления и хранения CID-идентификаторов.

Каждый CID-идентификатор состоит из трех частей: область совместимости, область типа и область информации. Две первых являются опциональными, но область информации обязательна. Область совместимости нужна для совместимости с существующими IoT ID сервисными схемами, состоит эта область из 8-битного кода страны/организации (COC) и 8-битного кода системы именовании (NSC).

Область типа нужна для реализации эффективного менеджмента и статистического анализа IoT-идентификаторов, различая схемы, объекты наименования и области применения разных систем наименования. Область типа состоит из 4-битного типа кодирования (CT), который специфицирует объем области информации, 4-битного типа ресурсов (RT), который используется для спецификации типа IoT-ресурса, например, штрих-код, RFID, сенсор и так далее, и 8-битного типа бизнеса (BT), который специфицирует область применения IoT-идентификаторов, например, сельское хозяйство, производство и т.д.

Область информации необходима для спецификации детальной информации поименованных IoT-ресурсов, например, сущность ресурса, свойства и т.д.

Схемы идентификации ISO TC46

Стандарт TC46/SC 9 ISO (международные стандарты) включает в себя такие идентификаторы как: ISBN (номер книг), ISSN (серийный номер), ISRC (код записи), ISMN (номер музыки), ISAN (номер аудиовизуальных файлов), ISWC (код музыкальных произведений), ISTC (текстовый код), ISNI (идентификатор имени), ISCI (коллекция идентификаторов). Все эти стандарты

имеют определенный набор связанных с ними метаданных, причем каждый набор независим от других, и не существует общих моделей или словарей для этих данных. Многие из этих идентификаторов не могут быть выражены в стандартной форме URI.

ISAN (International Standard Audiovisual Number) (ISAN) является уникальным идентификатором для аудиовизуальных произведений и смежных вариантов, аналогичных ISBN для книг. Он был разработан в рамках рабочей группы ISO TC46 / SC9. Управлением схемой идентификации занимается организация ISAN-IA. Стандарт (ISO 15706: 2002 и ISO 15706-2) рекомендуется в качестве аудиовизуального идентификатора для производителей, студий, вещателей, интернет-провайдеров, СМИ, производителей видеоигр. Идентификатор представляет собой уникальный международный номер, являющийся ссылкой для каждого аудиовизуального произведения, зарегистрированного в системе ISAN¹.

ISMN (International Standard Music Number) – буквенно-цифровой идентификатор для нотных партитур, разработанный ISO, состоящий из тринадцати цифр².

ISWC (International Standard Musical Work Code) – аналог ISBN для музыкальных произведений³.

ISRC (International Standard Recording Code) – международный стандартный код уникальной идентификации звукозаписей и музыки видеозаписи. Код был разработан компаниями индустрии звукозаписи совместно с техническим комитетом № 46 ISO и подкомитетом №9 (TC 46 / SC 9). ISRC идентифицирует конкретную аудиозапись, таким образом, различные записи, редакции и ремиксы одной и той же работы должны иметь свой собственный ISRC⁴.

1 International Standard Audiovisual Number URL: <http://www.isan.org/> (дата обращения: 13.12.2015)

2 The International ISMN Agency URL: <http://www.ismn-international.org/> (дата обращения: 13.12.2015)

3 The International ISWC Agency URL: <http://www.iswc.org/> (дата обращения: 13.12.2015)

4 IFPI URL: http://www.ifpi.org/content/library/isrc_handbook.pdf#Heading321 (дата обращения: 13.12.2015)

ISSN (International Standard Serial Number) – восьмизначный серийный номер используется для однозначной идентификации в серийном издании. ISSN используются для заказа, каталогизации публикаций и межбиблиотечного абонемент⁵.

ISTC (International Standard Text Code) – идентификатор для текстовых работ, решает проблему идентификации соответствия содержания, когда произведения опубликованы под разными названиями, или там, где различное содержание публикуется под одним и тем же названием. Другое применение ISTC предполагает использование идентификатора для идентификации отдельных, но связанных работ. Например, библиографические записи для ряда переводов, которые автоматически группируются вместе⁶.

LCCN (Library of Congress Control Number) – серийная система нумерации каталогизации записей Библиотеки Конгресса Соединенных Штатов Америки⁷.

SICI (Serial Item and Contribution Identifier) – международный стандартный серийный номер, который идентифицирует всю серию, а не конкретную книгу⁸.

ISNI (International Standard Name Identifier) – международный стандартный идентификатор имени был разработан в качестве стандарта для «именования» публичных лиц, которые участвуют в медиа-индустрии в области создания, производства, управления и содержания сетей дистрибуции. ISNI предназначен для идентификации создателей, но не правообладателей.

ORCID (Open Researcher and Contributor ID) – это международная, междисциплинарная, открытая организация, включающая научно-исследовательские институты, финансирующие организации, издателей и исследователей для повышения

5 International Standard Serial Number URL: <http://www.issn.org/understanding-the-issn/standardization/> (дата обращения: 13.12.2015)

6 URL: <http://www.istc-international.org/> (дата обращения: 13.12.2015)

7 Library of Congress URL: <http://www.loc.gov/ala/an-2008-update.html> (дата обращения: 13.12.2015)

8 URL: http://www.niso.org/apps/group_public/project/details.php?project_id=75 (дата обращения: 13.12.2015)

эффективности научного процесса и улучшения взаимодействия и эффективности финансирования научных исследований. ORCID призвана решить проблему неопределенности имен в научной коммуникации, создав реестр устойчивых уникальных идентификаторов для исследователей и открытый и прозрачный механизм между ORCID, другими схемами идентификации и исследовательскими объектами, такими как публикации, гранты и патенты. Схема идентификации ORCID совместима с ISNI, хотя изначально считалось, что она будет альтернативой ISNI.

LEI (Legal Identity Identifier) – переводится как допустимый идентификатор идентичности. Сфера применения LEI – это институты, которые держат финансовые активы для сектора финансовых услуг. LEI – специализированный идентификатор, совместимый со схемой ISNI.

ISBN

Идентификатор ISBN, уникальный номер, позволяет идентифицировать любое книжное издание.⁹ Международный стандартный номер книги (ISBN) является уникальным числовым коммерческим книжным идентификатором. Международный стандартный номер книги присваивается каждому изданию и его изменениям (за исключением переизданий) книги. Например, одна и та же книга как уникальная комбинация символов может существовать в виде электронной книги, в мягкой обложке и твердом переплете. Важно, *что эти формы изданий одной и той же книги будут иметь разные ISBN*. ISBN является последовательностью из 13 цифр, если назначен с 1 января 2007 года, и из 10 цифр, если назначен до 2007 года. Метод присвоения номера ISBN варьируется от страны к стране, часто в зависимости от того, насколько большой издательская индустрия является в данной стране.

Начальная конфигурация ISBN была создана в 1967 году на основе 9-значной Standard Book нумерации (SBN). SBN была создана в 1966 году, 10-значный формат ISBN был разработан

9 International ISBN Agency URL: <https://www.isbn-international.org/> (дата обращения: 13.12.2015)

Международной организацией по стандартизации (ISO) и был опубликован в 1970 году как международный. (SBN) код – числовой коммерческий идентификатор на основе девяти цифр, созданный Гордоном Фостером для книготорговцев и торговцев канцелярскими товарами WHSmith и других в 1965 году. Международное агентство ISBN работает в качестве регистрирующего органа для ISBN во всем мире, и ISBN стандарт разработан под контролем Технического комитета ISO / 46 Подкомитета 9 ТК 46 / SC 9.

SBN может быть преобразован в ISBN, предваряя цифру «0». Например, книга, опубликованная в 1965 году, имеет «SBN 340 01381 8» – 340 с указанием издателя, 01381 их серийный номер, а 8 контрольную цифру, это может быть преобразовано в ISBN 0-340-01381-8, контрольный символ не должен быть пересчитан.

Стандарт ISO 2108 (код SBN может быть преобразован в десять цифр ISBN с префиксом 0).¹⁰

Иногда книга может появиться без печатного идентификатора ISBN, если она напечатана в частном порядке или автор не следовал стандартной процедуре ISBN. Однако, ISBN может быть присвоен книге после ее издания, в момент, когда автор сочтет нужным пройти установленную процедуру. Выдача ISBN производится в конкретной стране. Номера ISBN выдаются регистрирующим органом ISBN, который отвечает за эту страну или территорию, независимо от языка публикации. Диапазоны номеров ISBN, приписанные к конкретной стране, основаны на издательском профиле соответствующей страны, и, таким образом, диапазоны будут меняться в зависимости от количества книг и размера издателей, которые являются активными. Некоторые регистрирующие органы ISBN основаны на национальных библиотеках или работают при министерствах культуры и, таким образом, могут получать прямое финансирование от правительства, чтобы осуществлять свои услуги. В других случаях регистрационная служба ISBN обеспечивается

10 Internet Archive WayBackMachine URL: <https://web.archive.org/web/20110430024722/http://www.informaticsdevelopmentinstitute.net/isbn.html> (дата обращения: 13.12.2015).

библиографическими данными поставщиков, которые не финансируются правительством.

Идентификатор ISBN содержит метаданные, по которым можно найти соответствующее издание в розничных сетях, библиотеках или в интернете. Одному и тому же произведению могут быть присвоены разные коды ISBN, такое происходит, когда книга была переработана, и вышло новое издание.

- страна происхождения или группа стран, объединенная языком издания; присваивается Международным агентством ISBN. Число цифр в идентификаторе группы зависит от объемов выпуска книжной продукции (может быть больше одной), например: 0 и 1 – группа англоязычных стран, 2 – франкоязычных, 3 – немецкий, 4 – японский, 5 – русскоязычные страны (некоторые страны бывшего СССР, Россия), 7 – китайский язык, 80 – Чехия и Словакия, 600 – Иран, 953 – Хорватия, 966 – Украина, 985 – Республика Беларусь, 9956 – Камерун, 99948 – Эритрея. В целом, группам присвоены номера 0—5, 7, 600—649, 80—94, 950—993, 9940—9989, и 99900—99999;

- код издательства; присваивается Национальным агентством ISBN, при этом учитывается количество изданий, которое издатель намерен выпустить в свет. Более крупным издателям присваивается более короткий номер, чтобы сделать доступным больше знаков для нумерации изданий (суммарная длина номеров издателя и издания для ISBN, присваиваемого российским агентством, составляет восемь цифр).

- уникальный номер издания (в России – от 6 до 1 знака);
- контрольная цифра (арабская от 0 до 9 или римская X); служит для проверки правильности числовой части ISBN. Расчет производит национальное агентство ISBN.

Особенности функционирования ISBN в разных странах:

В Канаде номера ISBN выдаются бесплатно с целью поощрения канадской культуры. В Великобритании, США и некоторых других странах, где предоставляется услуга неправительственными бюджетными организациями, выдача номера ISBN требует уплаты пошлины. Правительственным агентством, которое отвечает за выдачу номера ISBN, является Библиотека и Архив

Канады. Работы на французском языке ISBN идентифицирует по архивам Библиотеки Квебека.

В Австралии номера ISBN выдаются коммерческим агентством библиотечных услуг *Top-Bowker*, а цены варьируются от \$ 42 для одного ISBN (плюс \$ 55 регистрационный взнос для новых издателей) до \$ 2,890 для блока 1000 ISBN. Доступ осуществляется по немедленному запросу через веб-сайт.

В Бразилии за выдачу ISBN отвечает Национальная библиотека правительственного агентства, стоимость составляет примерно 16 долларов.

В Индии выдачей номера ведаёт Библиотечный фонд (RRRLF), являющийся подразделением Министерства культуры. Оно несет ответственность за регистрацию индийских издателей, авторов, университетов, институтов, государственных ведомств, которые несут ответственность за издание книг.

В Италии ответственность за выдачу ISBN несет организация *EDISER SRL*, принадлежащая *Associazione Italiana Editori* (Итальянской ассоциации издателей). Оригинальный национальный префикс 978-88 предназначен для издательских компаний, цена начинается с 49 евро 10-значный код. Существует также новый префикс 979-12, предназначенный для самостоятельной публикации авторов, по фиксированной цене 25 евро за один код.

Пакистан: Национальная библиотека Пакистана несет ответственность за регистрацию ISBN для пакистанских издателей, авторов, университетов, институтов, государственных ведомств, которые издают книги.

Южная Африка: Национальная библиотека Южной Африки несет ответственность за выдачу ISBN для южноафриканских издательских учреждений и авторов.

Великобритания и Ирландия: Частная компания *Nielsen Book Services Ltd*, часть *Nielsen Holdings NV*, отвечает за выдачу номера ISBN в блоках 10, 100 или 1000. Цены начинаются от 120 английских фунтов (плюс НДС) для наименьшего блока.

Соединенные Штаты Америки: В Соединенных Штатах деятельность осуществляется частной компанией ISBN. Существует

оплата, за изменения в зависимости от количества купленных номеров. Цена начинается от 125 долларов за один номер. Доступ осуществляется по немедленному запросу через веб-сайт.

Издатели и авторы в других странах могут получить номер ISBN из своих представительств национального агентства ISBN. Каталог агентств из ISBN доступен на сайте ISBN.

Метаданные, благодаря которым происходит идентификация книжного издания, были подвержены стандартизации, последняя из которых завершилась в 2007 году. Стандартизация необходима для создания удобного и понятного каталога, который будет храниться в репозитории. Стандартизация подразумевает создание единого алгоритма присвоения идентификационного кода из 13 цифр. Поводом для пересмотра алгоритма присвоения ISBN послужил дефицит числовых комбинаций. В настоящее время штрих-коды существуют по стандарту EAN-13; они могут иметь отдельный кодирующий штрих-код из пяти цифр для валюты и рекомендованной розничной цены. 10-значный номер ISBN, число «978», «код страны», является приставкой к ISBN в данных штрих-кодов, и контрольная цифра пересчитывается по формуле EAN13 (по модулю 10, 1 и 3-кратным взвешиванием на альтернативных цифрах).

Вследствие ожидаемого дефицита номеров в некоторых категориях ISBN и перехода ЕС на стандарт штрих-кода EAN-13, Международная организация по стандартизации (ISO) в 2007 году решила перейти на тринадцать цифр ISBN (ISBN-13). В 2011 году все 13-разрядные номера ISBN начинались с цифровой комбинации (префикса) 978. В ISBN был введен 979 префикс. Часть 979 префикса зарезервирована для использования для музыкальных партитур с ISMN.

Совместимость форматов штрих-кодов сохраняется, потому что формат штрих-кода ISBN-13 идентичен формату штрих-кодов EAN существующих 10-значных ISBN. Подвижность в системе EAN позволяет книготорговцам использовать единую систему нумерации для обеих книг и не-книжной продукции, которая совместима с существующими данными на ISBN основе, только с минимальными изменениями в области информационных

систем. Многие книготорговцы (например, Barnes & Noble) изменили в EAN штрих-коды уже в марте 2005 года.

Идентификаторы типа ISBN ориентированы на более специализированный поиск. Осуществляя запрос, пользователь получает доступ к базе данных, которая ему необходима, а не ко всему ассортименту информационных хранилищ, в которых упомянут искомый объект. При использовании такого идентификатора пользователь получает узкоспециализированную информацию об объекте, в то время как DOI предоставляет более расширенную информацию в результатах поиска.

ECODE

Entity Code (Код объекта) – это система идентификации для интернета вещей, которая предложена организацией Article Numbering Center of China. Каждый ECODE-идентификатор состоит из трех разных частей, включая версию (V), идентификатор системы нумерации (NSI) и мастер данных (MD). Версия размером 4 бита используется для различения кодовых структур системы ECODE.

Идентификатор системы нумерации (NSI) указывает на код другой идентификационной системы. Следуя различению в версии, размер NSI может быть 8 бит двоичных чисел, 4 десятичных или 5 десятичных. Мастер данных используется для спецификации идентификационных кодов отрасли или прикладной системы. Он находится в ведении и поддерживается местной организацией управления каждой системой идентификации, включая структуру кодирования и принцип распределения идентификаторов.

OID

Object Identifier (идентификатор объекта) является общей схемой кодирования, рекомендованной для однозначной идентификации объекта в глобальном масштабе. OID имеет очень хорошую основу для глобального применения. Сегодня OID успешно используется во многих сферах, таких как информационная безопасность, управление сетями, сенсорные сети, службы электронного здравоохранения и т.д. Более 100 IoT компаний зарегистрированы в OID через министерства, комитеты,

предприятия и научно-исследовательские институты. Китайская рабочая группа по сенсорным сетям (WGSN) поддерживает стандартизацию OID.

Структурно идентификаторы OID состоят из иерархической системы узлов, имена в которой задаются в соответствии со стандартом МСЭ-Т ASN.1 X.690. X.690 – один из стандартов ASN.1, разработанных совместно организациями ISO, ИЕС и МСЭ-Т для удобства представления данных при их передаче в телекоммуникационных сетях. Правила кодирования, описанные в X.690, служат для представления структур данных, описанных по правилам ASN.1, в виде последовательностей байт. Такие последовательности удобнее передавать по линиям связи или сохранять в файлы, чем делать те же операции с самими структурами.

Базовые правила кодирования или BER – набор правил, объясняющий, как представить любую структуру данных, описанную согласно ASN.1, в виде последовательности 8-битных октетов. Для того, чтобы различные типы данных можно было описывать схожим образом, в X.690 была определена общая структура блока закодированной информации, состоящая из следующих трех частей:

- Идентификатор – один или несколько октетов, в которых содержится информация о типе закодированных данных;
- Часть, содержащая информацию о длине блока – один или несколько октетов, в которых содержится информация о длине закодированных данных;
- Часть, содержащая закодированную информацию.

Прочие идентификационные системы

ASIN (Amazon Standard Identification Number) является 10-символьным буквенно-цифровым уникальным идентификатором, присвоенным Amazon.com и его партнерами по продукции идентификации в рамках организации Amazon.com.

Идентификаторы ASIN реализованы в нескольких местных региональных изданиях, в том числе в изданиях Великобритании, Франции, Германии, Италии, Канады и Японии. Несмотря на то, что ASIN претендует на экспансию в глобальном масштабе,

результаты поиска по номеру ASIN гарантированы только уникальным идентификатором в пределах рынка. Тот же самый продукт может называться несколькими ASIN, хотя и разные национальные сайты могут использовать разные ASIN для одного и того же продукта.¹¹

CODEN (идентификатор серийного издания, в настоящее время используется библиотеками; заменяется ISSN для новых произведений) в соответствии со стандартной E250 ASTM – это шесть символов, буквенно-цифровой библиографический код, который обеспечивает четкую, уникальную и однозначную идентификацию названий периодических изданий из всех предметных областей. Использование CODEN растет в научном сообществе как система цитирования для периодических технических и химических изданий, в качестве поискового инструмента во многих библиографических каталогах.¹²

ESTC (English Short Title Catalogue) является каталогом работ коллекции Британской библиотеки, опубликованных между 1473 и 1800 годами, в основном в Великобритании и Северной Америке, в первую очередь на английском языке.¹³

ETTN (Electronic Textbook Track Number) является уникальным цифровым идентификатором электронных книг, журналов и конференций. ETTN основан на 13-значном коде, созданном ассоциацией Magnanimitas Assn. ETTN коды могут быть использованы только для электронных текстовых файлов. Любой издатель, университет, ученый или профессор может запросить уникальный код для своего текста в электронном формате. ETTN состоит из 5 частей – AAA-BB-KKKKK-DD-F. AAA. Первая 3-значная последовательность кодируется согласно статистической классификации видов экономической деятельности в Европейском сообществе (КДЕС) и отображает тематику текста. Последовательность BB (2-значный) есть собственно код

11 Amazon URL: <http://www.amazon.ca/gp/help/customer/display.html?nodeId=200576730> (дата обращения: 13.12.2015)

12 Cassi URL: <http://cassi.cas.org/> (дата обращения: 13.12.2015).

13 British Library URL: http://estc.bl.uk/F/?func=file&file_name=login-bl-estc&local_base=bll06 (дата обращения: 13.12.2015)

ETTN. KKKKK (5-значный) – специальный идентификационный код для текста. ДД (2-значный) – месяц генерации текста. F (1-значный) – контрольная цифра.

Модели идентификации цифровых объектов на основе децентрализованного распределенного реестра

Децентрализованные реестры, цельность которых не зависит от центральной базы данных, пришли в мир с изобретением технологии блокчейна, самое известное применение которой – криптовалюта биткоин. Такой реестр основан на схеме, где центры доверия и управления передаются в виртуальную сеть, узлы которой постоянно записывают транзакции в определенном порядке, в общедоступные блоки, создавая тем самым цепочку (блокчейн).

Блок транзакций или блокчейн – специальная структура для записи группы транзакций в системе. Чтобы транзакция считалась достоверной («подтвержденной»), ее формат и подписи необходимо проверить, а затем уже группу транзакций записать в специальную структуру – блок. Информацию в блоках можно быстро перепроверить, а каждый блок всегда содержит информацию о предыдущем блоке. Все блоки можно выстроить в одну цепочку, которая содержит информацию о всех совершенных когда-либо операциях с блоками.

Блок состоит из заголовка и списка транзакций. Заголовок блока включает в себя свой хеш, хеш предыдущего блока, хеши транзакций и дополнительную служебную информацию. Первой транзакцией в блоке всегда указывается получение комиссии, которая станет наградой пользователю за созданный блок.

Далее идут все или некоторые из последних транзакций, которые еще не были записаны в предыдущие блоки. Для записи транзакций в блоке используется древовидное хеширование, аналогичное формированию хеш-суммы для файла в протоколе BitTorrent. Транзакции, кроме начисления комиссии за создание блока, содержат внутри атрибута input ссылку на

транзакцию, по которой на данный счет были получены биткойны. Комиссионные транзакции могут содержать в атрибуте любую информацию (для них это поле носит название `Coinbase parameter`), поскольку у них нет родительских транзакций.

Созданный блок будет принят остальными пользователями, если числовое значение хеша заголовка равно или ниже определенного числа, величина которого периодически корректируется. Так как результат хеширования (функции SHA-256) непредсказуем, не существует алгоритма получения желаемого результата, кроме случайного перебора. Если хеш не удовлетворяет условию, то произвольно изменяется блок служебной информации в заголовке, и хеш пересчитывается. Обычно требуется достаточно большое количество пересчетов и, соответственно, вычислительных мощностей. Когда вариант найден, узел рассылает полученный блок другим подключенным узлам, которые проверяют блок. Если ошибок нет, то блок считается добавленным в цепочку и следующий блок должен включить в себя его хеш.

Величина целевого числа, с которым сравнивается хеш, корректируется через каждые 2016 блоков. Теоретически заложено, что вся сеть будет тратить на генерацию одного блока примерно 10 минут, на 2016 блоков – около двух недель. Если 2016 блоков сформированы быстрее, то цель немного уменьшается и достичь ее становится труднее, в противном случае цель увеличивается. Изменение сложности вычислений не влияет на надежность сети и требуется лишь для того, чтобы система генерировала блоки почти с постоянной скоростью, не зависящей от мощности сети.

Каждый блок представляет собой контейнер с данными, доступ к которому может получить только владелец контейнера, но процедуру аутентификации владельца может провести любой узел сети. В случае биткойна хранимые данные представляют собой суммы валюты, однако технология может быть использована для хранения и идентификации любых цифровых объектов. Для построения подобных систем используются так называемые умные контракты: небольшие программы, записываемые вместе с блоком данных, которые определяют правила,

по которым эти данные будут использоваться. Основная идея умных контрактов в том, что стороны могут самостоятельно верифицировать операции, договорившись об условиях. Таким образом, метаданные, в том числе сведения о владельце объекта, могут быть записаны внутри блоков, а блокчейн, ко всему прочему, отвечать за систему резолюции.

В последние годы технология блокчейна находится в зените гартнеровского цикла известности (Gartner Hype Cycle), так что неудивительно, что появляется множество проектов, использующих ее для организации доверенных вычислений, идентификации и аутентификации объектов. В настоящий момент существует несколько решений, находящихся на различных стадиях разработки, от концептуальной модели до внедрения.

Namescoin представляет собой блокчейн с минимальными изменениями относительно биткойна, но внутри блоков в системе могут содержаться любые данные размером в пределах 520 байт. Основное предназначение системы – хранение информации о доменных именах .BIT, но создатели подчеркивают, что она может использоваться для хранения и резолюции любых подходящих идентификаторов.

Ethereum – универсальный блокчейн для реализации умных контрактов. В отличие от более простых решений на основе данной технологии, в блоках хранятся не только данные, но и определенные правила смены владения объектов, формата метаданных или содержания объектов. В отличие от многих криптовалют, токены EТН платформы Эфириум, называемые эфиром, позиционируются авторами не как альтернатива фиатным валютам, а как некое виртуальное топливо для учета использования ресурсов децентрализованной сети. В настоящее время токены продаются на биржах криптовалют, а капитализация всех токенов EТН достигла одного миллиарда долларов. Умный контракт Эфириума представляет собой, по сути, внутренний скрипт (обладающий полнотой по Тьюрингу), знающий текущее состояние системы. Таким образом, Ethereum позволяет реализовать сложные схемы множественной резолюции объектов.

Open Ledger Project – проект под управлением Linux Foundation с участием крупнейших финансовых корпораций и ИТ-компаний (Accenture, Cisco, IC3, IBM, Intel, J.P. Morgan, LSE, SWIFT, VMware, Wells Fargo и другие) Open Ledger вырос из концепции IBM Adept – автономной децентрализованной пиринговой телеметрии, блокчейна для идентификации и управления интернетом вещей. Open Ledger разрабатывает ПО для создания любых систем на основе децентрализованного реестра, не привязанных к определенному блокчейну.

ADEPT – еще одна идея, порожденная в рамках совместного проекта IBM и Samsung, радикально расширяет представление о перспективах глобального применения решений на основе технологии блокчейна. Крупнейший мировой производитель программного и аппаратного обеспечения и ИТ-сервисов объединил усилия с южнокорейским высокотехнологичным гигантом для совместного продвижения услуг и решений на рынке интернета вещей (IoT). Весной 2015 г. на сайте IBM был опубликован документ, описывающий концепцию новой технологической платформы – ADEPT (Autonomous Decentralized Peer-To-Peer Telemetry), ориентированной для продвижения услуг IoT. Архитектурная основа платформы предполагает интеграцию трех программных решений.

TeleHash – децентрализованный и защищенный пиринговый (P2P) протокол для обмена данными и передачи сообщений, первый прототип которого возник в 2010 г., а в данный момент используется спецификация конца 2014 г. В рамках концепции безопасности, данные и сообщения, передаваемые с использованием протокола, подтверждаются и заверяются третьей стороной; при этом модель коммуникации является одноранговой, модель «клиент-сервер» не используется.

BitTorrent – популярный пиринговый (P2P) сетевой протокол для кооперативного обмена файлами, созданный еще в 2001 г. и реализовавший концепцию обмена файлами через взаимодействие клиентов-источников (seeders and leeches).

Третьим компонентом является Ethereum – основанная на блокчейне виртуальная машина и набор сервисов Web 3.0,

дающий пользователям возможность работать с программной средой умных контрактов, развивая и наполняя ее контентом по их усмотрению за счет поддержки контрактного программирования.

Используя сочетание этих программных решений и платформ, IBM и Samsung планируют разработать глобальную экосистему умных объектов. Входящие в нее умные устройства смогут, имея доступ в интернет, обмениваться между собой данными и формировать единую глобальную непрерывную цепочку записей-транзакций, по аналогии с блокчейном биткоина. Отличие блокчейна ADEPT состоит в том, что содержание и типы записей в ней будут определяться теми контрактами, которые будут заключать между собой узлы сети. В случае со стиральной машиной, например, первый контракт будет сформирован при ее покупке и будет включать в себя данные о производителе, гарантийном обслуживании, технических характеристиках. Далее, в процессе эксплуатации устройство сможет само формировать и исполнять контракты на закупку порошка (при условии, что магазин порошка также будет подключен к ADEPT), диагностировать у себя поломку, определять, подпадает ли она под условия гарантийного ремонта и формировать заказ на такой ремонт отдельным контрактом с вендором, обращаясь к узлу гарантийной службы в ADEPT, и проч. Записи всех транзакций будут храниться в рамках единого глобального дерева, подобно общему блокчейну транзакций с биткоином. Понимая, что каждый узел ADEPT в каждом простом устройстве типа пылесоса, телевизора или смартфона не сможет хранить всю копию глобального дерева контрактов, разработчики системы ввели три уровня пиров с разным объемом ресурсов памяти, набором функций в рамках ADEPT и уровнем доступа к полной копии дерева контрактов. «Слабые» и ограниченные по функциям пиры в пылесосах и стиральных машинах смогут обращаться к более ресурсообеспеченным узлам ADEPT и запрашивать необходимую информацию у них, включая верификацию своих контрактов в глобальном блокчейне.

Block Verify – стартап, использующий модифицированный блокчейн биткоина для борьбы с подделками. Пилотный проект

в Великобритании по идентификации лекарственных препаратов на различных участках логистической цепочки и в местах продаж реализован совместно с фармацевтическими компаниями.

Несмотря на интересные возможности, которые открывает использование систем на основе децентрализованного реестра, все текущие проекты, использующие блокчейн для идентификации объектов, находятся на очень ранних стадиях внедрения. При этом необходимо отметить отсутствие единых стандартов (или необходимость адаптации существующих) среди множества концепций. Технология блокчейна позволяет обойтись без централизованных посредников и обладает уникальной масштабируемостью, но расплатой за эти свойства является низкая производительность и значительные вычислительные мощности, требующиеся для поддержания транзакций. Вполне вероятно, что оптимальной схемой использования блокчейна будет его встраивание в существующую схему идентификации объектов в качестве альтернативной системы резолюции объектов или использование в качестве дополнительного центра доверия, например, при резолюции в системе возвращаемые метаданные могут содержать ссылку на соответствующий блок в децентрализованном реестре.

Системы DOA/DOI

История и текущее состояние

В середине 1980-ых годов американская Корпорация национальных исследовательских инициатив (CNRI) разработала инновационный подход к управлению информацией в интернете, который впоследствии эволюционировал в архитектуру цифровых объектов (Digital Object Architecture).

В контексте этой архитектуры, цифровой объект – это структурированные данные, не зависящие от платформы, которым присвоен постоянный уникальный идентификатор – хэндл (handle). Система резолюции связывает эти идентификаторы с информацией о состоянии цифровых объектов. Подобная информация может содержать местонахождение данного объекта

в интернете (т.е. URL) или требования к доступу и использованию объекта, информацию об аутентификации и т.п. Создатель объекта или авторизованный администратор предоставляет эту информацию с использованием инфраструктуры публичных ключей, которая интегрирована в DOA. Технология публичных ключей предполагает использование двух ключей для шифрования – публичного и частного. Архитектура DO (и технология реестров объектов) включает в себя инфраструктуру для распространения публичных ключей. Система резолюции поддерживает верификацию пользователей и верификацию реестра для целей контроля доступа, при условии, что в метаданных содержится информация для аутентификации. Кроме того, сами по себе коммуникации между пользователями и реестрами могут быть зашифрованы.

Таким образом, архитектура DO создана с целью безопасного управления информацией (публичной, частной или комбинированной) в течение очень долгого периода времени. Инфраструктурная составляющая архитектуры обеспечивает техническую интероперабельность для организаций, которые договорились об административной составляющей. При этом компоненты системы могут использоваться практически для любой формы управления информацией, будучи ограничены только аппаратными требованиями.

Базовая архитектура содержит три типа компонентов, количество которых внутри типа не ограничивается. Архитектура не зависит от технологий, лежащих в основе имплементации компонентов, так же как архитектура интернета не зависит от выбора сетевых компонентов. Другими словами, реализация архитектуры может расширяться, технологически эволюционировать, при этом хранимая информация будет оставаться доступной в будущем вне зависимости от изменений в технологиях.

Основные компоненты архитектуры таковы:

Репозиторий цифровых объектов (DO Repository), который используется для хранения и доступа к цифровым объектам. Количество репозиторий в системе не ограничено. Для помещения в репозиторий объекту необходимо присвоить уникальный

идентификатор. При этом цифровые объекты могут перемещаться между репозиториями через административный интерфейс с сохранением всех метаданных, в том числе прав доступа к объекту.

Система резолюции (Handle System) отображает известные идентификаторы в хэндлы, содержащие информацию о состоянии идентифицируемого объекта. Система поддерживает юникод и различные национальные языки. Чаще всего хэндлы содержат информацию о местонахождении объекта (физическом или логическом), условия использования объекта, информацию для подтверждения аутентичности объекта, публичные ключи шифрования и т.д. Двухуровневая система резолюции объектов состоит из глобального реестра (GHR) и локальных реестров (LHR). При установке локального сервиса (LHS) создается пара ключей, публичный отправляется в GHR, а сервису выделяется префикс и права на изменение записей. ПО Handle System доступно в интернете (<http://www.handle.net>), при этом создание локальных сервисов доступно для любой организации при условии регистрации и получения номера авторизации – префикса. Использование префиксов гарантирует отсутствие дублирования идентификаторов.

Реестр объектов (DO Registry) используется для определения коллекций цифровых объектов и регистрации объектов, которые могут располагаться в одном или нескольких открытых репозиториях, как определено администратором метаданных. Реестр также используется для извлечения сведений о ранее зарегистрированном объекте (место, свойства, авторы, обладатели прав и т.п.) по его идентификатору. Результатом поиска объекта по его идентификатору (или детальным сведениям о нем, если это объект материального мира) является набор сведений (карточек) из метаданных в человекочитаемом или машиночитаемом виде (через графический интерфейс или API), а также номер последней транзакции об изменениях в метаданных. Реестр использует репозитории для хранения собственных метаданных (как объект), что позволяет управлять доступом к объектам.

В последних версиях программного обеспечения реестры и репозитории объектов объединены: каждый репозиторий

содержит реестр собственных объектов, а каждый реестр использует репозиторий для хранения собственных метаданных. ПО доступно в интернете по свободной лицензии (<https://www.cordra.org/>).

Архитектура цифровых объектов может обеспечивать интероперабельность между различными системами управления идентификацией или информацией. Как и интернет, DOA не заменяет существующие информационные системы, а позволяет им работать между собой, безопасно обмениваться информацией и предоставляет решения для долгосрочного хранения информации в неизменном виде.

В целях реализации компонентов архитектуры некоторые группы заинтересованных сторон разработали структуры, полагающиеся на регистрационные агентства и членскую систему для администрирования системы идентификации объектов. Наиболее ярким примером такой системы является International DOI Foundation (IDF), которая была создана издательским сообществом в начале 2000-х, но распространилась и на другие предметные области.

Система идентификации и резолюции, используемая IDF и его регистрационными агентствами, Digital Object Identifier (DOI) является реализацией Handle System, разработанной CNRI. В 1997 году CNRI участвовала в проекте Министерства обороны США Advanced Distributed Learning по стандартизации и модернизации учебных и тренировочных курсов. Целью ADL было эффективное обеспечение персонализированного доступа к обучающим материалам, «в нужное время, в нужном месте». Для этого была создана цепочка реестров метаданных и репозитивов материалов, поддерживающая систему резолюции Handle и уникальные идентификаторы.

Система DOI стала результатом инициативы группы издателей в конце 1990-ых годов, распознавших необходимость уникальной и непротиворечивой идентификации контента, вместо того, чтобы ссылаться на него по его местоположению. International DOI Foundation была основана в 1998 году для разработки системы на основе существующих технологий и

стандартов (т.е. Handle System). Первое регистрационное агентство появилось в 2000 году. Наиболее известной сферой применения DOI является Crossref (торговая марка ассоциации PILA) – универсальный сервис цитирования, который позволяет ученым ссылаться в работах непосредственно на источник цитаты, вне зависимости от издателя. С момента основания IDF активно работал с проектом *indecs* по построению фреймворка отображения словарей (Vocabulary Mapping Framework, VMF). VMF – это система автоматической разметки метаданных согласно общепринятой терминологии.

В 2000 году синтаксис DOI стал американским стандартом ANSI/NISO Z39.84, а в 2010 – международным стандартом ISO 26324. DOI является зарегистрированной частью схемы URI (RFC 4452), кроме того, идентификаторы DOI можно записывать как URL, резолюция при этом будет использовать сервисы HTTP Proxy. Наконец, как говорилось ранее, DOI является реализацией Handle System (IETF RFC 3650, 3651, 3652).

Разработка DOI происходила несколькими параллельными потоками. В первоначальной имплементации в хэндлах содержалась единственная ссылка на URL объекта или информацию о нем. Следующая версия системы обладала возможностями множественной резолюции, где результат зависел от контекста запроса. Одновременно происходила и разработка дополнительных инструментов по управлению информацией в цифровой среде (проект *indecs*). DOI может использоваться для постоянной идентификации различных объектов – физических, цифровых, абстрактных. Все объекты DOI имеют уникальные идентификаторы и описываются метаданными согласно принятой модели.

Модель данных

Объекты в системе DOI могут принимать любую форму: идентификатор может быть назначен любой сущности: физической, цифровой или абстрактной. Объекты в архитектуре DOA описываются как цифровые. Между двумя подходами нет конфликта, так как любая сущность может рассматриваться с точки зрения ее цифрового представления.

Синтаксис составления идентификаторов DOA и DOI идентичен. Лимита на длину самого имени или элементов суффикса или префикса не существует. Имена могут содержать любые печатные Unicode-символы и не зависят от регистра. Комбинация уникального префикса, выданного конкретному регистратору, и уникального суффикса, выдаваемого регистратором, очевидно, уникальна сама по себе, что и позволяет децентрализованно регистрировать идентификаторы. Префикс состоит из двух компонентов, директории и регистратора, разделенных точкой (индикатор директории у имен DOI всегда равен 10). Регистратору могут быть выданы многочисленные префиксы. После назначения объекту имени, оно не может изменяться, даже если объект перестал существовать, при этом в процессе резолюции такого объекта должно выдаваться соответствующее сообщение.

Несмотря на то, что в именах не рекомендуется содержать значимые данные (а указывать их в метаданных), суффиксы зачастую используются для включения идентификатора из другой системы, используемой для идентификации объекта. Например, суффикс DOI 10.1038/issn.0028-0836 может использоваться для периодического издания с идентификатором ISSN 0028-0836. В соответствующем контексте, например, при отображении в браузере, имена отображаются в форме URL после адреса прокси-сервера, например, <http://dx.doi.org/10.1038/issn.0028-0836>.

Объект, связанный с именем, однозначно описывается метаданными, основанными на расширяемой модели данных; назначение имени требует от регистранта заполнения минимального набора метаданных, требуемых для того, чтобы отличить объект от других в системе. В системе DOI этот набор – ядро – содержит другие известные идентификаторы и имя (или наименование) объекта. Кроме того, эти данные должны позволять определить фундаментальный тип объекта (физический, цифровой, абстрактный, пространственно-временной) и различать между объектами с одним и тем же именем (и одним создателем), но разными типами. Набор метаданных может расширяться регистрационными агентствами включением элементов, необходимых для работы соответствующих приложений.

Описание метаданных системы DOI¹⁴

Описательные элементы метаданных

Элемент	Число вхождений	Описание
DOI name	1	Идентификатор DOI
referentIdentifier(s)	0-n	Идентификаторы из других систем идентификации (ISAN, ISBN, ISRC, ISSN, ISTC, ISNI...). Для некоторых категорий идентифицируемых объектов это открытый список.
referentName(s)	0-n	Наименование идентифицируемого объекта. В этом элементе присутствует указание на язык согласно стандарту ISO 639-2 и на категорию в зависимости от природы идентифицируемого объекта.
primaryReferentType	1	Основная категория объекта. Открытый список. Среди стандартных допустимых значений: creation, party, event.
structuralType	1	Структурная категория объекта. Доступные варианты зависят от значения основной категории объекта. Для стандартных категорий эти списки ограничены. Так, для категории creation возможны 4 значения: physical, digital, performance, abstraction. Для категории party допустимы значения: person, animal, organization.
Mode	0-n	Этот атрибут представляет собой закрытый список форм восприятия, на которые рассчитан идентифицируемый объект. Атрибут предназначен только для категории creation. Допустимые значения: audio, visual, tangible, olfactory, tasteable, none.

¹⁴ Здесь приводится сокращенное описание ядра метаданных системы DOI. Формальная XML-схема, описывающая метаданные, доступна по адресу: http://www.doi.org/doi_schemas/DOIMetadataKernel_v2.2_131015.xsd. Словарь допустимых значений для атрибутов доступен по адресу: http://www.doi.org/doi_schemas/DOIAVS.xsd. Описываемые атрибуты имеют семантическое отражение в XML-схеме, но не совпадают с ними по наименованиям.

character	0-n	Этот атрибут представляет собой закрытый список форм, в которых выражен контент идентифицируемого объекта. Атрибут предназначен только для категории creation. Допустимые значения: music, language, image, other.
referentType	0-n	Роль по отношению к связанному объекту. Это открытый список со значениями, зависящими от основной категории объекта. Так, для типа party возможны, например, следующие значения, представленные атрибутом associatedPartyRole: Composer, Author, BookPublisher, JournalPublisher. Для типа creation обычно указывается формат, в котором представлен объект, описываемый атрибутом creationType. Например, возможны значения: audio file, scientific journal, musical composition, dataset, serial article, eBook, PDF.
linkedCreation	0-n	Производные от оригинального объекты. Этот элемент предназначен только для категории creation. В этом элементе могут быть указан атрибут creationRoleToCreation, характеризующий отношение производного объекта к оригинальному (например, экранизация книги).
linkedParty	0-n	Это элемент, применяемый только для категории party, и описывающий производные объекты категории party. В этом элементе может быть указан атрибут partyRoleToParty, уточняющий категорию производного элемента по отношению к исходному.
principalAgent	0-n	Это элемент, применяемый только для категории creation. Он предназначен для перечисления сторон, связанных с созданием или публикацией объекта. Это открытый список. Этот элемент содержит атрибут agentRole, специфицирующий конкретную роль стороны по отношению к объекту (например: Creator, Author, BookPublisher).

dateOfBirth-OrFormation	0-1	Это элемент, применяемый только для категории party и содержащий дату рождения (для человека или животного) или создания (для организации).
dateOfDeath-OrDissolution	0-1	Это элемент, применяемый только для категории party и содержащий дату смерти (для человека или животного) или прекращения существования (для организации).
associatedTerritory	0-n	Это элемент, применяемый только для категории party, и содержащий список территорий, с которыми ассоциирован данный объект, согласно стандарту ISO 3166a2.

Административные элементы метаданных

Элемент	Число вхождений	Описание
registrationAuthority-Code	1	Код, обозначающий присвоившее код RA.
issueDate	1	Дата присвоения данного DOI-кода.
issueNumber	0-1	Идентификатор используемой версии схемы метаданных DOI.

Система DOI не была предназначена для архивации идентифицированных объектов: в ней не хранятся сами по себе объекты, равно как и исчерпывающий набор метаданных, а только лишь указатели на источники подобных данных. При этом система позволяет обмениваться информацией об идентифицированных объектах, сохраняя при этом интероперабельность. Семантическая интероперабельность возможна благодаря модели данных DOI, состоящей из словарей данных и схемы их применения.

Принцип формирования и присвоения имен DOI

Как уже было отмечено, имя DOI является идентификатором (но не местоположения) объекта в цифровых сетях. Оно обеспечивает возможность для постоянной и доступной идентификации, а также взаимодействующего обмена управляемой

информацией в цифровых сетях. Имя DOI может быть присвоено любому объекту – физическому, цифровому или абстрактному – главным образом, для совместного использования с заинтересованной группой пользователей или управления им в качестве объекта интеллектуальной собственности. Система DOI разработана для обеспечения взаимодействия участников сетевых коммуникаций, а именно для использования идентификаторов и метаданных или работы с уже существующими системами.

Синтаксис имен DOI определяет создание скрытых строк с указанием владельцев имени и делегирования. Информация, содержащаяся в DOI электронного документа, содержит указатель его местонахождения (например, URL), его имя (название), прочие идентификаторы объекта (например, ISBN для электронного образа книги) и ассоциированный с объектом набор описывающих его данных (метаданных) в структурированном и расширяемом виде.

Каждое имя DOI представляет собой уникальную строку букв и цифр, состоящую из двух обязательных компонентов: префикс и суффикс, которые в совокупности образуют имя DOI, разделенные символом «/». Так, часть, следующая после разделительного символа «/» (суффикс), может быть действующим идентификатором или уникальной строкой, выбранной регистратором. А часть, которая находится до разделительного символа «/» (префикс), относится к уникальному владельцу имени и присваивается регистратором. Любая организация вправе иметь несколько префиксов.

Суффикс определяет метаданные об объекте, например, место первого издания, год издания, номер выпуска и порядковый номер статьи в выпуске (присваивается издателем но технически контролируется системой распознавания DOI).

Например:

10.1000/123abc,

где

10.1000 – префикс, или идентификатор издателя, составленный из признака идентификатора (10) и строки, указывающей на издателя (1000). Префикс присваивается регистратором DOI.

123abc – суффикс, идентификатор объекта, указывающий на конкретный объект. Суффикс формируется издателем, и должен быть уникальным у данного издателя.

DOI также способен обеспечить «контейнер» идентификатора, которой способен разместить любой существующий идентификатор (например, такие как ISBN, ISSN, ISTC) и др. При этом суффикс может формироваться по конструкции, принятой для того или иного идентификатора, или даже объединять их.

Пример использования в имени DOI суффикса с идентификатором ISSN:

10.1038/issn.0028-0836

Пример расширенной версии идентификатора:

10.1000/1111-2222-2015-1-33-44,

где

1111-2222 – ISSN публикации;

2015 – год издания;

1 – номер выпуска;

33-44 – номера страниц статьи в выпуске.

Цифровой идентификатор объекта регистронезависим.

К примеру, такие имена DOI как 10.1000/123abc, 10.1000/123ABC, 10.1000/123AbC являются идентичными.

При формировании суффикса допустимо использовать только следующие символы:

- буквы латинского алфавита: a-z, A-Z;
- цифры: 0-9;
- знаки препинания: – . _ ; () /

При отображении на экране или в печати имени DOI, как правило, предшествует обозначение “doi:” в нижнем регистре, по аналогии отображения таких сетевых протоколов (например, как “ftp:” или “http:”), если необходимо однозначно указать на принадлежность идентификатора к системе DOI.

Пример:

doi:10.1000/123abc

Ограничений по длине имени DOI не существует.

Имя DOI может быть присвоено любому объекту, который, в свою очередь, должен быть точно определен с помощью структурированных метаданных. Само имя DOI остается постоянным, невзирая даже на изменения прав собственности на него. Даже если контент будет видоизменен с течением времени, присвоенный ему идентификатор DOI останется неизменным.

Проект *indecs*

Понимание метаданных зависит от контекста. Каждое значение атрибута должно быть определено и идентифицировано так, чтобы учесть любой возможный контекст. В различных национальных языках существует множество синонимов и омонимов, значение слов может меняться в зависимости от профессиональной сферы, источника, представления и т.д. Для взаимопонимания между участниками системы необходимы подробные словари стандартных терминов (мы используем их во многих областях, для этого существуют такие организации, как ISO), которые всегда будут определяться однозначно. Основой для схемы метаданных и расширений, использующихся в системе DOI, является проект *indecs* (интероперабельность данных в системах электронной коммерции) – контекстный онтологический подход к описанию метаданных.

Проект *indecs* был создан в конце 1990-х годов европейским сообществом создателей и издателей контента для анализа требований к метаданным для коммерческого использования контента в интернете. В основу была положена модель жизненного цикла любого типа контента или интеллектуальной собственности, от задумки до физического или цифрового воплощения. В качестве главного принципа работы была предложена гипотеза, что можно создать механизм для обработки сложных метаданных различных типов контента таким образом, чтобы произведения воспринимались как объекты с различными значениями абстрактного набора атрибутов, общего для всех типов контента. При этом интероперабельность должна поддерживаться вне зависимости от медиума или языка.

Проект *indecs* предлагает четыре основных принципа управления идентификацией объектов:

- Принцип уникальной идентификации: каждая сущность должна быть уникально идентифицирована в идентифицированном же пространстве имен;
- Принцип функциональной детальности: в случае выделения новой сущности должна иметься возможность ее идентификации;
- Принцип определения прав: автор метаданных должен быть идентифицирован в безопасной манере;
- Принцип уместного доступа: каждому требуется доступ к метаданным, от которых они зависят, и приватность и конфиденциальность их собственных метаданных от тех, кто от них не зависит.

Фреймворк *indecs* в дальнейшем развивался в направлении общего онтологического подхода к определению различных типов сущностей и атрибутов и их связей в зависимости от контекста. Этот фреймворк используется в следующих приложениях и проектах:

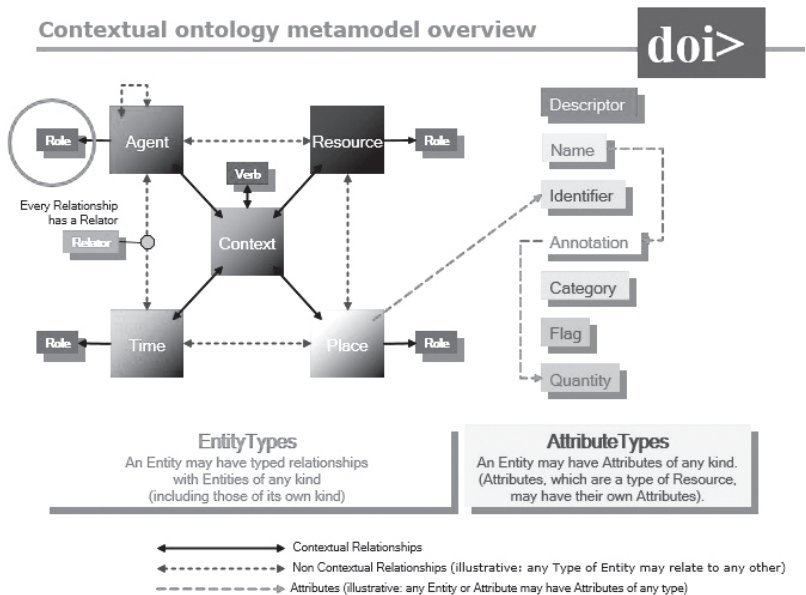


Рисунок 2. Онтологическая модель данных DOI

- Linked Content Coalition
- RDA/ONIX Framework for Resource Categorization
- DDEX (Digital Data Exchange, проект музыкальных издателей)
- ONIX (Online Information Exchange, проект книжных издателей)
- Vocabulary Mapping Framework (общественный проект, в котором активно участвует фонд IDF). VMF – это справочник стандартизированных метаданных и инструментов для автоматической разметки метаданных по наиболее подходящей схеме из принятых.

VMF предоставляет возможность автоматизации разметки метаданных, упрощая использование информации в различных контекстах. То есть словари терминов являются по сути репозиторием для всех элементов данных, атрибутов и возможных значений, которые они могут принимать, использующихся в спецификации метаданных системы.

В рамках эволюции архитектуры цифровых объектов в настоящий момент идет разработка так называемых реестров типов (Type Registries), которые предназначены для создания пользователями собственных семантических структур для описания объектов. Объекты реестра типов имеют уникальный разрешимый идентификатор, который возвращает характеристику структур, соглашений, семантику и представление данных. Такой объект упрощает людям и машинам понимание и обработку данных. Таким образом, в метаданных цифрового объекта DOA/DOI можно указать ссылки на соответствующие сервисы или приложения (которым также присвоен идентификатор в реестре типов). А в метаданных приложения – объекта реестра типов – будет храниться API этого приложения.

Прототипы реестра типов создаются в настоящий момент многими научными и государственными организациями, в частности:

- Национальный институт стандартизации и технологий (NIST), США
- Обсерватория глубокого углерода (DCO), США
- Инициатива генома материалов (MGI), США
- Американское бюро переписи.

Реестр и репозиторий объектов

Одними из главных элементов архитектуры DOA являются реестры и репозитории объектов. Первоначально программное обеспечение для создания реестров было разработано CNRI в качестве дополнения к системе резолюции Handle для использования в системе DOI. Однако, с эволюцией архитектуры DOA программное обеспечение Cordra стало включать в себя как реестр, так и репозиторий цифровых объектов, то есть теперь каждый репозиторий содержит реестр собственных объектов, а каждый реестр использует репозиторий для хранения метаданных. Число репозиториев не ограничено, а метаданные могут легко перемещаться из одного в другой, при этом сам репозиторий может быть представлен в виде цифрового объекта в системе.

Основные функции Cordra:

- 1) регистрация цифрового объекта;
- 2) извлечение сведений об объекте (место, свойства, авторы, обладатели прав и т.п.) ранее зарегистрированного цифрового объекта по его идентификатору;
- 3) поиск объекта или детальных сведений о нем (если объект материального мира);
- 4) получение номера последней транзакции, сведений об изменениях метаданных;
- 5) хранение информации об объектах и обеспечение доступа к ним.

Технология реестра может быть использована для построения распределенной системы источников данных с различной архитектурой. Наиболее простым подходом было бы создание единого реестра, куда регистраторы могли бы передавать метаданные о сущностях для проверки, выдачи идентификатора и индексации. Роль реестра можно свести лишь к хранению минимально необходимых данных, тогда как регистратор может поддерживать дополнительные сервисы (в том числе нестандартные) хранения данных. Или же система резолюции может хранить ядро данных, а реестр – расширенные данные, а роль регистратора сводиться только к отправке сведений в систему. Может существовать схема с множеством параллельных реестров, изменения в которых бы зеркалировались.

Таким образом, существующие программные решения могут использоваться для создания альтернативной архитектуры и не обязательно должны использоваться в рамках DOA.

Handle System

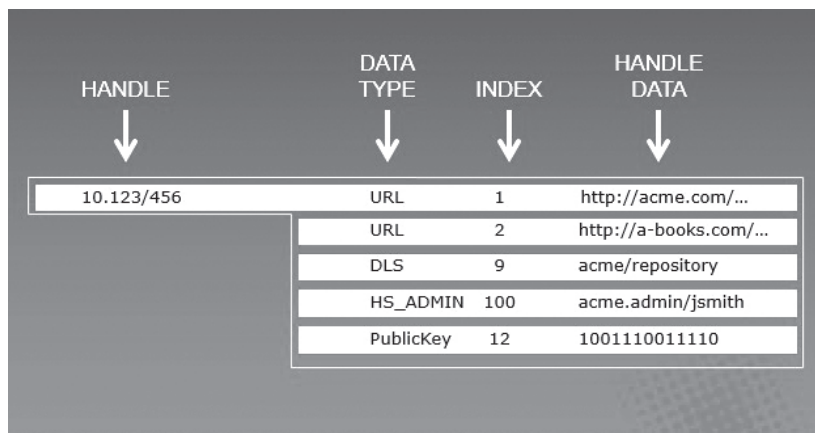


Рисунок 3. Множественная резолюция хэндлов

Система резолюции (разрешения; хэндл; Handle system) была создана, чтобы преодолеть ограничения функциональности существующих систем идентификации объектов в интернете. Сам вопрос о выборе системы для идентификации объектов возник в конце 1990-х. Были сформулированы базовые требования к будущей системе идентификации. Разработка системы была произведена CNRI совместно с DARPA. Основными качествами, заложенными при ее создании, были:

1. Уникальность – каждый хэндл уникален в рамках глобальной системы;
2. Постоянство – хэндлы могут использоваться в качестве постоянных идентификаторов для объектов в интернете. При этом хэндл не зависит от объекта, который он именуется, их единственная связь – в самой системе. Это качество не гарантирует постоянство само по себе, так как является задачей администрирования. Но оно позволяет имени существовать неизменным вне зависимости от изменений местоположения, владения и т.д. То есть при перемещении ресурса достаточно

обновить его значение в Handle System для отражения нового местоположения;

3. Множественные экземпляры – хэндл может указывать на различные экземпляры ресурса, расположенные по различным сетевым адресам. Приложения могут использовать это качество системы для повышения производительности и устойчивости;

4. Множественные атрибуты – хэндл может указывать на различные атрибуты ресурса, включая связанные сервисы, расположенные по различным сетевым адресам;

5. Расширяемое пространство имен – локальное пространство имен можно присоединить к глобальному, получив статус регистратора и уникальный префикс, чтобы избежать конфликтов с существующими именами. Использование регистраторов позволяет делегировать сервисы администрирования и резолюции локальным сервисам (local handle service), которым будет передавать запросы глобальный реестр (Global Handle Registry), создавая распределенную модель;

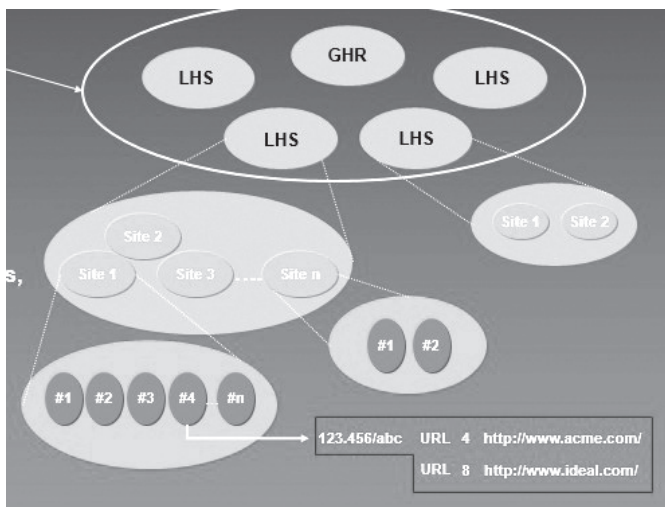


Рисунок 4. Распределенная модель Handle System

6. Модель безопасности – Handle System позволяет осуществлять безопасную резолюцию и администрирование. Протокол системы определяет стандартные механизмы клиентской и

серверной аутентификации и авторизации, а также содержит функции проверки целостности данных и ограничения приватности;

7. Распределенный административный сервис – для каждого хэндла в системе можно определить собственного администратора (владельца). В комбинации с протоколом аутентификации это позволяет администратору безопасно управлять хэндлом через интернет;

8. Эффективный сервис резолюции – протокол системы спроектирован высокопроизводительным. Так, например, интерфейс администрирования отделен от резолюции и не влияет на его производительность. При этом система позволяет осуществлять чрезвычайно тонкое разделение функций административного контроля, в отличие от других распределенных систем, где, как правило, существует небольшое количество администраторов с широкими полномочиями.

Резолюция – это процесс, в котором идентификатор является запросом к сетевому сервису на получение актуальной информации (данных о состоянии), относящихся к определяемой сущности, чаще всего – местоположение. Handle System поддерживает множественную резолюцию, то есть ответом на запрос может быть местоположение различных экземпляров объекта, связанные сервисы, и любая другая информация, указанная в метаданных объекта. Процедура резолюции может возвращать все возможные варианты, а может только конкретный, в соответствии с запросом клиента и типом приложения. В текущей реализации DOI объекты со схожей структурой метаданных могут быть сгруппированы в профили приложений; в свою очередь профили приложений могут быть связаны с различными сервисами. Таким образом, сервис можно связать с множеством идентификаторов, добавив его к профилю приложений. В будущем это решение будет заменено на реестр типов объектов.

Возвращаемая информация не обязательно должна указывать собственно на экземпляр объекта: например, это может быть описание или состояние объекта, некие индикаторы или измерения, отношения с другими сущностями и т.д. Процесс выбора

результата при множественной резолуции автоматизирован и незаметен для пользователя.

Для обработки запроса через протокол http система использует прокси-серверы, которые обрабатывают хэндлы, представленные в виде URL. Варианты резолуции при этом хранятся в хэндл-записи в виде xml-файла. За работу прокси также отвечает корпорация CNRI, как за необходимую составляющую поддержания целостности миллионов веб-ссылок на хэндлы объектов. Таким образом, прокси-серверы, например, doi.org, являются надстройкой над Handle System.

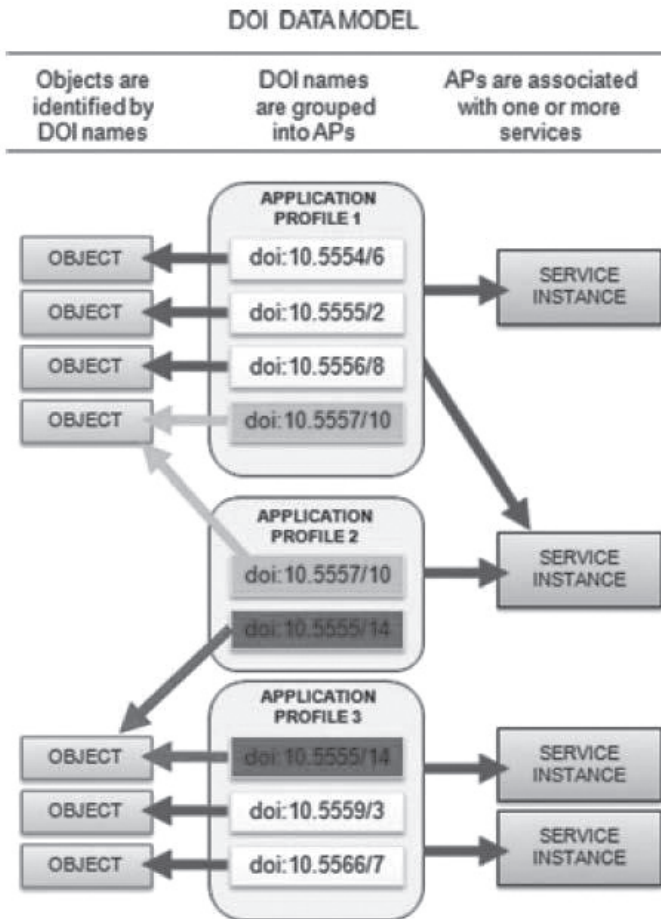


Рисунок 5. Множественная резолуция и профили приложений

Handle System описана в RFC 3650-3652:

- RFC 3650, “Handle System Overview”;
- RFC 3651, “Handle System Namespace and Service Definition“;
- RFC 3652, “ Handle System Protocol (ver 2.1) Specification”.

Handle System предусматривает использование идентификаторов, состоящих из двух частей: префикса и суффикса. По префиксу можно однозначно установить организацию, занятую его обслуживанием (администратора). Соответствие префикса его администратору хранится в Global Handle Registry (GHR). Суффикс однозначно идентифицирует конкретный объект. Информация о соответствии суффикса объекту хранится в Local Handle Registry (LHR).

Предусмотрен двухэтапный резолвинг каждого идентификатора. Приложение-клиент сначала запрашивает в GHR информацию о резолвере конкретного префикса, получая в ответ информацию о LHR, а затем делает запрос к LHR уже с полным идентификатором, получая в ответ набор данных об интересующем объекте. Протокол предусматривает использование стандартных технических решений, таких как кеширование, наличие прокси-серверов и т.п.

Каждому объекту в Handle System приписан ряд обязательных атрибутов:

1) Числовой идентификатор данных в наборе (индекс). Это позволяет, в частности, хранить в одной записи об объекте исторические данные, разноплановые представления информации и т.п.;

2) Тип данных. Тип данных представляет собой ссылку на запись в Handle-системе. В RFC 3651 определен ряд базовых типов данных (URL, администратор и др.). Возможно использование других типов данных, определенных на уровне конкретной модели данных;

3) Время последнего изменения данных;

4) TTL (Time To Live);

5) Права доступа для администратора записи и для неавторизованного пользователя. Для описания прав используется модель, близкая к модели Unix (права на чтение, запись и исполнение);

- 6) Ссылка на администратора данной записи об объекте;
- 7) Ссылки на другие handle;
- 8) Данные. Запись в этом атрибуте зависит от того, какой тип данных указан для объекта.

Эта система является расширяемой: так как в рамках каждого префикса информацию об объекте отдает администратор префикса, то он может добавлять типы данных, характерные для его предметной области, и осуществлять резолвинг самостоятельно, отталкиваясь от предложенной им самим семантики.

В протоколе Handle System Protocol, специфицированном в RFC 3652, предусмотрен ряд возможностей для уточнения получаемого набора данных, в частности:

- выбор типа возвращаемых данных;
- выбор индекса или диапазона индексов;
- указание на то, можно ли ответить закешированными данными.

В качестве одной из мер защиты в протоколе применяется электронная подпись возвращаемых данных, что придает ответу достоверность и неотрицаемость. По умолчанию ответы серверов Handle не подписаны, необходимость подписи должна быть явным образом указана в запросе. Необходимость явного требования подписи позволяет экономить вычислительные ресурсы для высоконагруженных серверов.

Преимущества системы Handle перед прочими системами адресации

Вопрос о преимуществах системы Handle перед другими системами адресации/идентификации подробно рассматривается в разделе 6 RFC 3650.

Преимуществом системы Handle по сравнению с системой DNS является более гибкая модель администрирования префиксов и лучшая масштабируемость при увеличении количества суффиксов в выбранном префиксе.

Преимуществом системы Handle перед системой Directory Services (LDAP, X.500) является ограниченность базовой функциональности, сводящейся для Handle к резолвингу заранее

известного указателя (имени), что обеспечивает более высокую устойчивость к нагрузкам. При этом системы, построенные на базе Handle, могут предоставлять возможности для поиска, аналогичные LDAP-системам.

Система Handle, в отличие от большинства прочих систем адресации/идентификации, обеспечивает персистентность идентификаторов. В отличие от системы URL/URN, в которой, по разным оценкам, 10-20% ссылок становятся невалидными в течение года (т.н. “rotten links”), идентификация средствами системы Handle обеспечивает владельцам идентифицируемых объектов возможность актуализировать ссылки и метаданные, поддерживая их в актуальном состоянии.

В отличие от публикаций в Web, где для различения оригинальной публикации от перепечатки может потребоваться нетривиальная экспертиза, в случае с системой Handle аффилированные с объектом акторы заинтересованы в поддержании информации в актуальном состоянии, а также в том, чтобы объект был уникально идентифицирован (например, из соображений цитируемости – в случае научных статей).

Распространенность и применение системы Handle

На базе системы Handle построены многие системы регистрации объектов, как цифровых (например, D-Space, <http://dspace.org>), так и физических (система проверки подлинности продуктов, построенная в Китае, позволяет идентифицировать отдельные упаковки товаров).

Система Handle позволяет развивать произвольные сервисы, основанные на идентификации с помощью идентификаторов Handle. Общее наименование такого рода сервисов – DOA, Digital Object Architecture.

Наиболее популярным публичным явлением, основанным на системе Handle, является группа проектов (система), объединенных именем DOI (Digital object identifier, цифровая идентификация объектов). С точки зрения адресации система DOI располагает первичным префиксом 10 и порожденными префиксами. По состоянию на конец 2015 года в системе DOI было зарегистрировано более 15000 префиксов и более 114

миллионов объектов. В то же время в прочих системах было зарегистрировано около 3000 префиксов. Точное количество объектов, зарегистрированных в прочих системах, неизвестно. По некоторым оценкам, приводящимся в интернете, количество объектов, зарегистрированных вне рамок DOI, превышает 1 миллиард.

Безопасность Handle System и взаимодействие с инфраструктурой публичных ключей (PKI)

Как мы уже упоминали ранее, обеспечение гибких возможностей для администрирования, защита персональных данных и конфиденциальности, обеспечение целостности данных и безотказность являются требованиями к системе идентификации цифровых объектов в интернете. Handle System на уровне архитектуры обеспечивает эти функции безопасности, в том числе путем аутентификации клиента и сервера. Handle System поддерживает инфраструктуру публичного ключа (PKI), причем для ее развертывания не нужны дополнительные сервисы – подобная информация может храниться непосредственно в метаданных объектов. Впрочем, при необходимости для аутентификации могут использоваться и сторонние сервисы, а также системы симметричного шифрования.

При развертывании локального сервиса (LHS) и регистрации в системе Handle, для локального сервиса создается пара ключей, публичный ключ при этом отправляется в глобальный реестр (GHR), а сервису выделяется префикс и права на изменение записей. Все идентификаторы, связанные с локальным сервисом, будут подписаны его ключом, таким образом, сторонний сервер не может изменять данные, ассоциированные с другими префиксами, и подобные попытки будут пресекаться на уровне GHR. В свою очередь вся информация, передающаяся из GHR в локальный сервис, будет подписана публичным ключом, что позволяет поддерживать целостность системы и не дает злоумышленникам подделать запросы от глобального реестра.

Для обеспечения целостности данных клиент может установить зашифрованное соединение с серверами системы и запросить цифровую подпись данных; возвращаемые сервером

данные могут быть при необходимости зашифрованы перед отправкой клиенту. Для дополнительной гарантии целостности также могут использоваться сторонние сервисы, например, возвращаемые метаданные могут содержать ссылку на цифровую подпись из доверенного источника.

По умолчанию, данные, хранимые в системе, доступны публично, однако каждый хэндл может иметь определенные ограничения к просмотру, а содержимое может быть зашифровано. Каждый хэндл в системе может иметь собственного администратора, с определенным набором привилегий. Протокол Handle аутентифицирует администратора, прежде чем выполнить любой административный запрос или резолюцию данных, для которых администратор задал определенный уровень конфиденциальности. По умолчанию, процесс резолюции не требует аутентификации клиента, но в случае наличия ограничений ответ будет выдан клиенту только после проверки публичного ключа. То есть таким клиентам (и администраторам) присваивается собственный хэндл и ассоциируется с релевантной информацией о пользователе и персональным сертификатом для безопасности. Для повышения производительности процессы выполнения административных запросов и запросов о резолюции разделены в системе.

Прокси-серверы, обеспечивающие трансляцию запросов к хэндлам через протокол http, не являются частью системы Handle, то есть ответы от прокси-серверов не могут аутентифицироваться через систему. Таким образом, администраторы должны отдельно озаботиться их защитой от злоумышленников, в особенности от атак типа man-in-the-middle.

Идентификация объектов в контексте интернета вещей

Интернет вещей предполагает миллиарды ежедневно используемых взаимосвязанных устройств, предоставляющих такие возможности по созданию новых приложений, которые не были бы возможны без комбинации всех этих «умных» объектов.

Вопросы идентификации и адресации в такой огромной сети становятся ключевыми для работы других аспектов системы. По сути интернет вещей – это использование информационных технологий для осуществления взаимосвязи уникально идентифицируемых объектов. IoT-приложения обычно основаны на координации информации и сервисов различных сетевых объектов, как физических (например, сенсоры), так и виртуальных или логических (например, ПО «умного дома», сервис электроснабжения). Таким образом, наличие механизмов идентификации для всех типов объектов является ключевым условием для разработки, внедрения и работы нетривиальных решений на основе интернета вещей.

Идентификация вещей (IDoT) – это область задач по присвоению уникальных идентификаторов и связанных метаданных объектам интернета вещей, позволяющая им обмениваться информацией с другими сущностями в интернете. При этом «вещью» может быть любая сущность, имеющая идентификатор и возможность передачи данных: и физическая, и логическая. Однако, в отличие от классического управления идентификацией, в рамках интернета вещей необходимо учитывать некоторые особенности обеспечения безопасности, а именно:

Жизненный цикл. Некоторые объекты IoT могут иметь долгий жизненный цикл. Электронная медицинская запись, к примеру, это логический объект, который сохраняет *идентичность* в течение всей жизни пациента. С другой стороны, идентификатор посылки будет существовать только в течение процесса доставки почты.

Взаимоотношения. Важно понимать, как IoT-объекты связаны с другими сущностями, не входящими в интернет вещей – владельцами, администраторами, производителями, пользователями. Владелец, пользователь или администратор устройства может меняться с течением времени, что влияет на процессы идентификации, аутентификации и авторизации.

Знание контекста. Процессы управления идентификацией и доступом должны учитывать контекст. Например, возможна такая ситуация, при которой необходимо разрешить

пользователю доступ к системе, но в другой ситуации открыть доступ для того же пользователя будет небезопасным (скажем, в умной системе оповещения о пожаре).

Механизмы защиты. Методы подтверждения идентичности в классическом управлении идентификацией складывались в течение многих лет. Аутентификация, проверка целостности, отказоустойчивость встроены в такие протоколы как SAML или OpenID. В интернете вещей же протоколы связи часто построены не на основе интернет-протоколов, а ресурсы устройств весьма ограничены. Если устройство использует всего несколько байт для передачи данных, тут не может быть места для шифрования или других механизмов безопасности.

Аутентификация. Мультифакторная аутентификация эффективна для применения в отношении людей, но не устройств, так как многие факторы являются биометрическими, а пароль в M2M-коммуникациях заменяется токеном или сертификатом. Мультифакторная аутентификация осуществляется комбинацией следующих доказательств идентичности: «Что-то, что у тебя есть»; «Что-то, что ты знаешь»; «Что-то, чем ты являешься». Последние два пункта неприменимы для интернета вещей.

Множественные или уникальные идентификаторы? В случае появления глобальной схемы идентификации, по аналогии с IP, необходимо прояснить вопросы управления критической инфраструктурой для такой схемы. Новые стандарты протоколов на основе IPv6, такие как 6LoWPAN, показывают, что можно создать эффективную схему присвоения уникальных идентификаторов для очень малых устройств. Но для разрешения вопросов непересекающихся адресных пространств в глобальном масштабе требуется инфраструктура, поддерживающая динамичные устройства, постоянно появляющиеся или исчезающие из сети и перемещающиеся между различными публичными и закрытыми сетями; поддерживающая идентификацию и аутентификацию пользователей и защиту персональных данных; позволяющая получать и обмениваться информацией об объектах и их метаданными. При создании такой схемы необходимо учитывать, что во многих отраслях экономики устоялись собственные

схемы идентификации, часто закрытые или проприетарные, и переход на полностью новую систему может быть не оправдан экономически и политически.

Идентификаторы или сетевые адреса? Идентификаторы служат уникальным указателем на объект, тогда как сетевой адрес может меняться в зависимости от физического местоположения или логического участия в той или иной сети. В тех случаях, когда идентификатор и адрес различаются, для идентификации объектов используются более подходящие схемы, такие как, например, EPC для RFID-тэгов или его японский аналог uCode. Точно так же могут существовать различные схемы адресации. Для объектов в интернете используется IPv4 или IPv6, но объекты могут использовать и другие частные схемы и протоколы. При этом они могут быть доступны через интернет благодаря специальным шлюзам, транслирующим данные между глобальной и локальной сетью.

Поиск и резолюция. Поиск объектов является тривиальной задачей для маленьких сетей, но требует значительных ресурсов с увеличением числа объектов. При этом в динамической среде интернета вещей автоматизированные механизмы обнаружения необходимы для взаимодействия между объектами и сервисами. Система DNS используется не только для трансляции доменов в IP-адреса и наоборот, но и, например, почтовыми агентами для поиска места доставки сообщений – общий механизм для поиска сервисов в домене с использованием SRV-записей и NAPTR-записей для резолюции объектов. Для интернета вещей существуют расширения системы DNS Object Naming Service (ONS) и Object Directory Service (ODS), а также система Handle, которые используются для разрешения идентификаторов объектов.

Прозрачность и независимость от сети. В компьютерных сетях идентификаторы могут содержать информацию о местоположении устройств. Объекты интернета вещей должны иметь идентификаторы, не зависящие от того, в какой сети они находятся или каким пользователям принадлежат.

Расширяемость до миллиардов устройств. Один из наиболее изученных методов создания схемы идентификации с

поддержкой миллиардов объектов – это использование иерархического наименования в зависимости от контекста, местоположения или домена для устранения возможных конфликтов имен.

Эффективность для очень простых устройств. Объекты интернета вещей могут иметь очень ограниченные вычислительные способности (сенсоры), или же их может не существовать вовсе (RFID-тэги). Схемы идентификации должны работать даже в таких условиях, когда, например, устройству доступно всего несколько килобайт памяти.

Защита персональных данных. IoT-устройства будут собирать огромное количество персональной информации. Схемы идентификации должны поддерживать различные уровни доступа к данным для различных пользователей и защищать собираемые данные по умолчанию.

Гибкость и расширяемость. Исходя из предполагаемого количества объектов, схема идентификации для интернета вещей должна быть гибкой и расширяемой. Например, должна существовать возможность легко обновлять местоположение объекта или добавлять новые объекты к определенным группам устройств. Надежность и интероперабельность также являются важными качествами: в идеале любой пользователь должен иметь возможность быстро подключить приобретенное устройство, не отвлекаясь на технические детали (что в настоящий момент происходит весьма редко). При этом аутентификация и идентификация, скорее всего, будут первыми шагами, которые должен будет предпринять пользователь при подключении устройства.

Безопасность. Большинство современных решений для идентификации IoT фокусируются на предоставлении масштабируемого сервиса именования и адресации, при этом безопасность находится на втором плане. Обязательный функционал системы должен включать в себя аутентификацию при доступе к данным в процессе поиска или резолюции; авторизацию и проверку прав при доступе к данным об именовании и адресации, проверку таких данных на подлинность; шифрование данных при обмене

и предотвращение перехвата пакетов; предотвращение манипуляций с кэшем системы; защиту от DoS-атак; предотвращение регистрации зловредных объектов. Стоит еще раз отметить, что все перечисленные особенности системы идентификации должны создаваться с учетом масштабируемости: процедура авторизации при большом количестве объектов может быть весьма ресурсоемкой задачей.

В настоящий момент для управления идентификацией вещей используется набор различных технологий, однако эти идентификаторы можно разделить на несколько уровней и для целей использования.

- Идентификаторы объектов – используются для идентификации физических или виртуальных объектов. (EPC, UPC, Handle/DOI, UUID, MAC, Ecode, OID, CID)

- Коммуникационные идентификаторы – используются для идентификации устройств в процессе обмена данными через сеть. (IPv6)

- Идентификаторы приложений – используются для идентификации сервисов и приложений в рамках интернета вещей. (URL, URI)

Работа над стандартизацией и разработкой систем идентификации и управления доступом для интернета вещей ведется на различных площадках, например, в рамках экспертной комиссии Еврокомиссии и Китая; IETF работает над модификацией существующих IoT-протоколов в рамках проекта Аутентификации и авторизации для ограниченной среды (ACE). Так, протокол Delegated CoAP protocol позволяет делегировать процессы аутентификации более ресурсным узлам сети, снижая аппаратные требования для устройств. Вместе с тем, на текущий момент сложно сказать, приведут ли усилия по стандартизации к глобальному единому стандарту, или же появятся несколько различных схем идентификации с разным уровнем совместимости.

Термин «интернет вещей» используется в контексте миллиардов различных устройств, но в реальности это, в первую очередь, управление информацией об этих устройствах. И главным отличием от привычного интернета является устройство конкретных

сервисов, в частности, идентификации. В настоящий момент существует несколько примеров внедрения систем для именованя, адресации и обнаружения объектов интернета вещей (IPv6, Handle, EPC/ONS и др.). Однако преждевременно говорить, что какая-то из этих систем станет доминирующей, скорее они будут существовать одновременно, обслуживая различные запросы, области применения и географии. Большинство из них было создано не для интернета вещей, а для расширения возможностей адресации в интернете или отслеживания цифровых или физических объектов. В результате они не являются универсальными, а для успешной работы масштабных приложений требуются решения, поддерживающие интероперабельность между различными технологиями идентификации. В свете ограничений существующей инфраструктуры интернета и протоколов, построение систем управления информацией интернета вещей на их основе выглядит неразумным.

В то же время, открытая архитектура интернета до сих пор позволяла развивать инновации в коммуникациях и приложениях без внесения фундаментальных изменений в базовые протоколы. Среди существующих и апробированных технологий управления информацией, разделяющих подобный подход, можно выделить архитектуру DOA и ее ключевую часть – Handle System, которая существует уже более 15 лет. Гибкий механизм резолюции идентификаторов дает возможность для эволюции системы и сервисов на ее основе. При этом Handle System является распределенной системой, функции которой распределены между множеством локальных провайдеров. Со временем, вполне вероятно и появление новых систем, обладающих качествами DOA, например, на основе децентрализованного реестра.

Раздел 2.

Реестр цифровых объектов как инструмент идентификации контента в цифровую эпоху

Особенности правового регулирования гражданско-правового оборота объектов авторских прав в сети Интернет

Стремительное развитие информационных технологий и глобальной сети Интернет радикально меняет подходы к распространению информации, создавая дилемму для старых индустрий и порождая множество вопросов по поводу эффективных механизмов контроля над распространением информации. В последнее время очевидно, что и публичные органы власти, осуществляющие охрану общественных интересов в связи с распространением противоправной информации, и корпоративные правообладатели исключительных авторских/смежных прав на произведения ставят перед собой совершенно разные цели, но они едины в своем стремлении контролировать частные коммуникации граждан в глобальной информационной среде.

Наибольшую сложность при современном уровне развития децентрализованных сетей и криптографии представляет охрана исключительных прав на произведения в сети Интернет. За годы своего стремительного развития интернет, каким мы его знаем сегодня, включил в себя миллиарды различных произведений, каждое из которых имеет автора. Часть из них находится под охраной национального законодательства и международных конвенций, другая часть распространяется на условиях свободных лицензий, и еще одна – значительный объем контента

– является UGC-контентом¹⁵, образующим уникальные системы Web 2.0, которые путем учета сетевых взаимодействий становятся тем лучше, чем больше людей ими пользуются.

При этом правовой статус большого количества произведений, представленных в глобальной сети, можно обозначить как “сиротские”, то есть произведения, авторство которых не установлено, и правообладатели которых не выразили своей воли предоставить этот контент широкому кругу лиц в пользование, либо иным способом распорядиться принадлежащим правом на оговоренных автором условиях.

Несмотря на появление все более широких технических возможностей копирования цифровых объектов и доступа к самому разнообразному контенту, неавторизованное копирование материалов, защищенных авторским правом, представляет большую проблему для медиа-индустрии. Эту причину называют в качестве основной, мешающей реализации продаж развлекательного контента в сети. Стремясь снизить уровень сетевого пиратства, провозглашая высшей целью защиту прав создателя, многие страны пытаются сегодня различными способами модернизировать национальное и международное право, которое отстает от развития сетевых технологий. Такие попытки предпринимаются для того, чтобы закон отвечал современным вызовам традиционной модели копирайта, что позволило бы передовым игрокам медиаиндустрии эффективно справляться с неконтролируемым копированием информации в глобальном цифровом пространстве.

Воспитание легального потребления контента и наказание пользователей

Одной из первых стран, которая решила бороться с сетевым пиратством путем наказания конечных пользователей пиратского контента, стала Франция.

Принятый в 2009 году закон HADOPI¹⁶ установил создание специальной федеральной комиссии HADOPI, задачами

15 UGC – User generated content – контент, создаваемый самими пользователями сети (к примеру, ЖЖ, YouTube, Wiki)

16 <https://en.wikipedia.org/wiki/HADOPI> _law

которой являлось управление творческими произведениями и защита авторских прав в интернете, а также специальную процедуру общения с нарушителем авторских прав.

Так, после подачи заявления правообладателем и предоставления им IP-адресов нарушителей, HADOPI направляет пользователю информацию о том, что последний нарушает закон. Такое письмо направляется по e-mail через провайдера. Ему также предлагается установить фильтр для своего подключения к интернету. Если в течение 6 месяцев после этого правообладатель, его представитель, провайдер или HADOPI поймут пользователя еще на одном скачивании пиратского контента, ему направляется повторное письмо с требованием прекратить нарушение исключительных прав. В случае если в течение года пользователь совершит еще одно правонарушение, HADOPI вправе направить в суд заявление о взыскании с нарушителя штрафа в размере до 1500 евро, а также отключении ему интернета (с запретом всем другим провайдерам оказывать услуги связи проштрафившемуся пользователю).

Закон вызвал много споров и подвергся рассмотрению Конституционным советом, верховным судебным органом республики, и в итоге был признан нарушающим конституционные права граждан: во Франции, как и в других странах ЕС, доступ к интернету признан базовым правом любого человека.¹⁷

На практике, система трех предупреждений (“трех страйков”), после чего пользователь-рецидивист должен был быть отключен от интернета, оказалась нежизнеспособной. Первое предупреждение о нарушении авторских прав из-за скачивания или распространения пиратских файлов получили 2 миллиона французов. Второе предупреждение поступило 186 тысячам пользователей; всего состоялось 663 судебных заседания в связи с третьим нарушением, но репрессиям подвергся лишь 1 человек. Судом было предписано отключить его от

17 Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Frank La Rue // http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A.HRC.23.40_EN.pdf

сети и взыскать 600 евро штрафа. Но даже это наказание не было приведено в исполнение из-за решения правительства в 2013 г. отменить закон. Во-первых, решением Верховного Конституционного совета страны он был признан неконституционным. Во-вторых, он практически не работал.

Другим примером попытки воспитать пользователей и предупредить их о недопустимости нелегального копирования является антипиратская программа CAS (Copyright Alert System), реализованная в США в 2013 году в качестве корпоративного соглашения между представителями крупнейших медиакорпораций RIAA и МРАА, с одной стороны, и операторами связи Verizon, AT&T, Cablevision, Comcast, Time Warner и прочими, с другой стороны.

По идее авторов указанной программы, большинство пользователей не понимают, что, скачивая копию фильма с ближайшего торрент-трекера, они нарушают закон. Таким образом, до них необходимо донести, что участие в раздачах в p2p-сетях защищенного копирайтом контента является незаконным. Для этого интернет-пользователям, которые нарушают авторские права, направляются последовательно несколько предупреждений: от нейтральных до более строгих. Если нарушитель не отреагировал на 6 предупреждений, провайдер может замедлить скорость соединения или вовсе прекратить обслуживание клиента.

Основная идея CAS состояла в том, чтобы отказаться от насилия, которое до этого практиковалось в США в отношении интернет-пиратов, в пользу сравнительно мягкого наставления на истинный путь. Можно с уверенностью сказать, что это была попытка медиа-лоббистов оптимизировать затраты. По открытым сведениям, Американская ассоциация звукозаписи (RIAA) потратила с 2006 по 2008 год около 64 миллионов USD на оплату расходов юридических фирм, сопровождавших судебные процессы, связанные с защитой авторских прав¹⁸. За тот же период ассоциации удалось получить от нарушителей

18 <http://recordingindustryvspeople.blogspot.ru/>

менее 1,4 миллиона долларов в виде штрафов и при внесудебном урегулировании конфликтов. Таким образом, убытки правообладателей для защиты авторских прав в сети составили более 62 миллионов USD.

Несмотря на попытки США навязать более жесткую охрану авторских прав на международном уровне и показать высокий уровень пиратства в странах третьего мира, как это представляется согласно докладу № 301 Торгового представительства США, уровень неавторизованного копирования в США остается на очень высоком уровне. Об этом наглядно свидетельствует последнее исследование «культуры копирования», предпринятое Джо Караганисом и Леннартом Ренкема, согласно которому в США и Германии нелегальное копирование охраняемых авторским правом произведений практикует от половины до 2/3 всех интернет-пользователей¹⁹.

На сегодняшний день деятельность CAS не дала никаких результатов. Кроме того, цифры опубликованных деклараций за 2014 год показали, что работа Информационного центра по авторским правам в США, ответственного за CAS, обошлась в 3 млн. долларов за год. Эта цифра вызвала множество вопросов у правозащитников и общественных деятелей, т.к. получилось, что на отправку одного предупреждения о пиратстве тратится порядка 2 долларов налогоплательщиков²⁰. Расходы провайдеров оказались выше, чем сумма, которую они платят Информационному центру. Провайдеры за свой счет также тратятся на технические средства, необходимые для обработки «Предупреждений об авторском праве» и дополнительную поддержку клиентов.

В 2011 г. Новая Зеландия также представила свою новую модель борьбы с пиратством²¹. Было решено выносить предупреждения пользователям, пойманым на распространении

19 <http://piracy.americanassembly.org/wp-content/uploads/2011/11/MPEE-Russian-1.0.pdf>

20 <https://torrentfreak.com/six-strikes-anti-piracy-scheme-costs-3-million-per-year-140729/>

21 <http://3strikes.net.nz/>

материалов, защищенных авторскими правами. В случае, если эти меры становятся неэффективными, после третьего предупреждения пользователь привлекается к суду по делам о нарушении авторских прав.

Согласно новозеландской модели, интернет-провайдер самостоятельно должен следить за пользователями, уведомляя их о нарушении исключительных прав, а суд принимает решение о наложении штрафа и применении наказания в виде отключения от интернета.

Однако анализ правоприменения показал, что на деле система оказалась абсолютно неработоспособной и довольно убыточной.

На провайдеров доступа была возложена обязанность рассылать нарушителям авторских прав официальные уведомления по почте, а на правообладателей – оплачивать эти уведомления, из оценки 25\$/письмо. Общая сумма расходов по предъявлению одного обвинения за один случай нарушения копирайта составляет порядка 275\$.

Но на деле до судебного разбирательства доходит редко. Поэтому деньги правообладателей и работа провайдеров уходят впустую, а сетевые пираты в итоге не несут никакого наказания и не оплачивают штрафы. В 2014 году на нелегальном файлообмене было поймано всего 18 человек²².

В Великобритании под давлением правообладателей в 2015 году также была запущена кампания по рассылке пользователям предупреждений. При этом в адрес одного пользователя установлен лимит отправки: не более 4-х писем по обычной и электронной почте. Согласно корпоративному соглашению между правообладателями и операторами связи, с каждым письмом формулировки должны быть все жестче. ВРІ²³ и МРА²⁴ также требовали включения в письма пользователям правовых

22 <https://xakep.ru/2015/09/16/nz-3-strikes-law/>

23 British Phonographic industry – Ассоциация производителей фонограмм Великобритании

24 Motion Pictures Association of America – Ассоциация Кинокомпаний Америки

предупреждений и настаивали на получении доступа к базе данных пользователей, которые скачивают контент нелегально, чтобы иметь возможность в дальнейшем воздействовать на них в судебном порядке. Тем не менее, правозащитники и OFCOM²⁵ не согласовали возможность включения в письма угроз или намеков на последствия для нарушителей.

Сторонники жестких законодательных мер, которые предлагают наказывать за нелегальное копирование и дистрибуцию авторского контента, аргументируют необходимость принятия подобных мер тем, что цифровое пиратство серьезно бьет по доходам авторов и снижает их мотивацию на создание новых произведений. Однако подобная аргументация не выдерживает никакой критики.

В целом данные о производстве в творческих индустриях вполне согласуются с точкой зрения, согласно которой свободный файлообмен не демотивирует артистов и издателей. Несмотря на то, что в музыкальной сфере продажи альбомов несколько упали с 2000-х годов, число выпускаемых альбомов продемонстрировало взрывной рост. Так, например, в 2000 г. было выпущено 35516 альбомов. Семь лет спустя издано 79695 альбомов, включая 25159 цифровых альбомов. Даже если файлообмен был причиной падения продаж в целом, новая технология вовсе не стала погребальным звоном по количеству производимой в мире музыки. Подобные тенденции наблюдались и во многих иных творческих отраслях: музыке, программном обеспечении, компьютерных играх, фото, видео. Общее число производимого в мире контента неуклонно растет.

Блокировка сайтов и ограничение доступа к информации

В последнее время США являются главными инициаторами жесткого регулирования оборота интеллектуальной собственности. При этом в США существует две политики в области охраны результатов интеллектуальной деятельности: внутренняя, с принципом свободного (добросовестного) использования

25 OFCOM – The Office of Communications, уполномоченный орган в Великобритании в сфере связи, телекоммуникационных технологий, массовых коммуникаций и почты

объектов авторского права, иммунитетом для информационных посредников, невозможностью блокировки сайтов; и международная, направленная на максимальное ужесточение регулирования и игнорирование существующих международных структур, таких как ВОИС²⁶.

Первыми американскими законопроектами, предполагающими блокировку доступа к сайтам в сети Интернет, стали:

- Stop Online Piracy Act (*рус. – Законопроект о прекращении онлайн-пиратства*), известный как SOPA, который был внесен в Палату представителей США.
- Protect Intellectual Property Act (*рус. – Законопроект о защите интеллектуальной собственности*), известный как PIPA, который был внесен в Сенат США.

По сути, законопроекты представляли из себя переписанную версию ранее предложенного акта Combating Online Infringement and Counterfeits Act (COICA), который годом раньше не прошел чтения в Конгрессе США.

Оба закона были внесены в Конгресс США в 2011 году и содержали схожий порядок регулирования. По задумке авторов законопроектов, обновление законодательства об авторском праве в сторону расширения полномочий корпоративных правообладателей было экстренно необходимо для сохранения прибылей организаций, рабочих мест и защиты интеллектуальной собственности в соответствующих отраслях (фильмы, музыка, программное обеспечение и др.). Лоббисты от медиаиндустрии особенно отмечали то, что PIPA и SOPA помогут эффективно бороться с интернет-сервисами, расположенными за пределами США и вне их юрисдикции.

Однако с того момента, как тексты документов стали доступны интернет-общественности, оба они стали объектом непрекращающейся критики. Противники SOPA и PIPA утверждали, что законопроекты нарушают Первую поправку к Конституции США, вводят цензуру в интернете, ограничивают свободу слова и возможности развития сети Интернет.

26 Всемирная организация интеллектуальной собственности (World Intellectual Property Organization, WIPO)

Онлайн и офлайн протесты гражданских активистов, интернет-пользователей и крупных IT-компаний в 2012 году заставили конгрессменов и сенаторов оказаться от дальнейших планов работы над законопроектами, что означало большую победу интернет-сообщества и IT-бизнеса.

Несмотря на провал в США законопроектов, позволяющих блокировать веб-сайты в сети Интернет, попытки навязать единому цифровому пространству новую модель защиты авторских прав путем ограничения доступа к веб-ресурсам, нарушающим интеллектуальные права, не были оставлены, и вскоре подобные инициативы стали появляться в других национальных правовых системах и международных соглашениях.

Так, в Тайване был предложен к принятию законопроект, аналогичный скандальному американскому SOPA. Автором законопроекта стал уполномоченный федеральный орган по регулированию оборота интеллектуальной собственности (ТИРО). В соответствии с проектом ТИРО предлагало самого себя наделить полномочиями по ограничению доступа к иностранным сайтам, которые по единоличному решению ТИРО являются “ресурсами, предназначенными для массового нарушения авторских прав третьих лиц”. Законопроект был отозван в связи с серьезным противодействием интернет-сообщества и значительным освещением в СМИ.

В Сингапуре с 7 по 21 апреля 2014 года также были проведены общественные консультации по поводу реформы законодательства об авторском праве. В деле борьбы с пиратством сингапурские законодатели ушли еще дальше своих тайваньских коллег, приняв закон о внесении изменений в национальное законодательство об авторском праве²⁷.

Согласно указанному акту, национальные власти имеют право блокировать доступ к таким ресурсам как The Pirate Bay²⁸ или KickAssTorrent, грубо нарушающим законодательство об авторском праве. Закон вступил в силу с августа 2014 года и позволил

27 [http://www.parliament.gov.sg/sites/default/files/Copyright%20\(Amendment\)%20Bill%2016-2014.pdf](http://www.parliament.gov.sg/sites/default/files/Copyright%20(Amendment)%20Bill%2016-2014.pdf)

28 <http://ccm.net/news/25095-pirate-bay-and-more-blocked-in-singapore>

правообладателям подавать заявления в Верховный суд в целях дальнейшего направления предписания в адрес ISP об ограничении доступа к сайтам-нарушителям²⁹, как это следует из ст. 193А Закона о копирайте.

Так, в соответствии с новым законом Верховный суд Сингапура имеет право определять, используется ли тот или иной сайт в сети Интернет для постоянного грубого нарушения авторских прав третьих лиц. Суд делает выводы о деятельности сайта исходя из нескольких факторов:

а. является ли основной целью сайта нарушение авторских прав;

б. содержатся ли на сайте директории, индексы или категории произведений, нарушающих авторские права;

в. демонстрирует ли владелец или оператор сайта пренебрежение к охране исключительных авторских прав;

г. был ли ограничен доступ к сайту в соответствии с решениями судов других иностранных юрисдикций за нарушение исключительных авторских прав;

д. содержит ли сайт какие-либо инструкции и описание инструментов для восстановления доступа к сайту;

е. объем трафика на сайт.

Указанный список не является исчерпывающим, и суд может рассмотреть и другие условия для принятия мер по блокировке интернет-ресурса. В случае, если Верховный суд даст на то свое одобрение, правообладатели имеют возможность блокировать и удалять спорный контент или блокировать хостинг сайта в течение восьми недель.

Несмотря на то, что нормы антипиратского законодательства, предложенные в COICA, SOPA и PIPA, не прошли в Соединенных Штатах, многие положения одиозных законопроектов были экспортированы в рамках соглашения TPP (Trans-pacific partnership), что может в ближайшее время серьезно изменить ситуации с копирайтом для пользователей сети как в самих США, так и в странах, подписавших соглашение с США.

29 <https://torrentfreak.com/singapore-passes-pirate-bay-blocking-anti-piracy-law-140708/>

Киберполиция и уголовное преследование

Примером создания специализированного подразделения полиции, которое бы занималось правонарушениями в области интеллектуальных прав, является Великобритания. Так, в Соединенном королевстве существует PIPCU (Police Intellectual Property Crime Unity), которое представляет интересы FACT (Federation Against Copyright Theft). В полномочия указанного подразделения входит возможность административного делегирования доменных имен, рассылка штрафов за неавторизованную дистрибуцию и копирование материалов, защищенных копирайтом, а также привлечение пользователей и владельцев сайтов к уголовной ответственности по преступлениям, связанным с нарушением авторских прав.

В России функции киберполиции выполняет отдел “К” МВД России. В плотном сотрудничестве с Российской антипиратской организацией (РАПО), представляющей интересы крупнейших американских студий, сотрудники указанного подразделения уже неоднократно возбуждали уголовные дела против пользователей, уличенных в скачивании и раздаче материалов, защищенных авторским правом в файлообменных сетях.

Особенности правового регулирования копирайта в сети Интернет в России

Несмотря на провал антипиратского закона в США, в российское законодательство была осуществлена имплементация положений, аналогичных положениям скандальных законов SOPA и PIPA, предлагающих расширение правовых возможностей правообладателей и уполномоченных государственных органов.

Под давлением Торгового представительства США 21 декабря 2012 года США и Россия подписали двустороннее соглашение³⁰ об утверждении плана действий по совершенствованию охраны интеллектуальных прав, в котором предусматривалось ужесточение ответственности за нарушение авторских прав в сети, а

30 United States-Russian Federation Intellectual Property Rights Action Plan // <https://ustr.gov/sites/default/files/uploads/IPRAActionPlan%2014Dec12%20fin.pdf>

также предлагались иные методы для усиления охраны исключительных прав в российском сегменте интернета.

Через месяц после подписания межправительственного соглашения Минкульт РФ предложил первый проект «антипиратского закона», который частично копировал положения отвергнутых в США законопроектов SOPA и PIPA. Эта законодательная инициатива сразу же вызвала оживленную дискуссию в российском интернет-сообществе и IT-отрасли.

Невзирая на массовые протесты IT-компаний и пользователей, под напором группы иностранных и российских лоббистов от крупных правообладателей видео- и аудиоконтента Министерство культуры 6 июня 2013 года направило в Госдуму законопроект «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам защиты интеллектуальных прав в информационно-телекоммуникационных сетях».

26 июня 2013 г. РАЭК и Mail.ru Group, «Яндекс», Объединенная компания «АфишаРамблер-СУП», «Google Россия», RUCENTER, Хостинг-Центр, Фонд содействия развитию технологий и инфраструктуры Интернета, Wikimedia Russia, OZON.RU, Ассоциация интернет-издателей выступили³¹ с консолидированной позицией интернет-отрасли, объявив, что законопроект несет серьезные риски для бизнеса и общества.

Тем не менее, все замечания, подготовленные профильными министерствами и ведомствами, а также критика со стороны интернет-индустрии, были проигнорированы. В сжатые сроки законопроект, вводящий механизм ограничения доступа к сайту в связи с нарушением исключительных авторских прав на видеоконтент, был принят Федеральным Собранием РФ и уже 3 июля 2013 года подписан Президентом, став Федеральным законом № 187-ФЗ и получив неформальное название «антипиратский закон».

31 Открытое обращение интернет-индустрии в отношении законопроекта №292521-6 «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам защиты интеллектуальных прав в информационно телекоммуникационных сетях» // <http://raec.ru/times/detail/2667/#sel=>

В закон был заложен ряд весьма противоречивых положений, в том числе нормы, позволяющие заблокировать сайт не только за размещение на нем непосредственно самого объекта авторских прав, но и «информации, с помощью которой можно получить доступ к объекту, размещенному с нарушением исключительных прав правообладателя». Таким образом, доступ к любому сайту может быть ограничен даже за наличием на нем гиперссылки, торрент-файла, магнет-файла или иной информации, в т. ч. текстовой, о том, где можно скачать файл.

Несмотря на попытки самых влиятельных игроков IT-отрасли донести до лиц, принимающих решения, что закон вреден для российского интернета и несет большие риски для бизнеса, снижая при этом инвестиционную привлекательность российского сегмента мировой сети, закон был принят без учета каких-либо поправок и замечаний со стороны отрасли с вступлением его в силу с 1 августа 2013 г.

Широкая волна протестов и забастовок прокатилась по Рунету. 1 августа около 1600 сайтов выключили свои серверы или поставили протестные баннеры, в ряде городов прошли митинги. На следующий день после подписания закона на портале Российской общественной инициативы силами общественных организаций «Роскомсвобода», Пиратская партия России и Ассоциация пользователей интернета была запущена общественная петиция³² об отмене антипиратского закона. Петиция получила широкую поддержку среди пользовательских сообществ и операторов сайтов. В рекордно короткие сроки петицию подписало более 100 000 авторизованных пользователей сети, в связи с чем она приобрела статус обязательной к рассмотрению экспертной комиссией при Открытом правительстве. Однако 14 октября по итогам рассмотрения экспертной рабочей группы чиновники рекомендовали отклонить петицию.

На этом принятие дальнейших антипиратских мер не закончилось. Федеральный закон № 364-ФЗ, вступивший в силу

32 Общественная инициатива №50Ф4494 об отмене закона о произвольных блокировках интернет-ресурсов от 02.07.2013 № 187-ФЗ (закон против интернета) // <https://www.roi.ru/4494/>

с 1 мая 2015 года, расширил действие ранее принятого антипиратского закона на все объекты авторских прав (кроме фотографий), ввел дополнительные обязанности для владельцев веб-сайтов и предусмотрел еще большую ответственность для информационных посредников, установив ряд нововведений.

Закон обязывает указывать на сайте информацию о его владельце, о месте нахождения и адресах, в т.ч. адресе электронной почты для направления заявления от правообладателей, а также вводит возможность ограничения доступа к сайту навсегда за два и более нарушения исключительных прав третьих лиц. Кроме того, в законе так же, как и в первой версии, заложена возможность вечной блокировки интернет-ресурса не только за размещение на сайте самого произведения, но и за размещение информации, необходимой для его получения «без права снятия ограничения на доступ к сайту», что ставит под удар любые социальные сети и иные UGC-ориентированные веб-ресурсы, где не предусмотрена премодерация контента.

Очевидно, что законодательство об авторском праве не отвечает современным потребностям всех участников цифровых отношений (правообладателей, пользователей, информационных посредников) и не учитывает их интересы и права. При этом вместо принятия закона, аналогичного американскому варианту SOPA, существовал и альтернативный пример решения в виде ранее реализованного и ныне действующего американского закона в области оборота произведений, защищенных копирайтом, DMCA с более мягким регулированием и более эффективным взаимодействием стейкхолдеров.

Параллельно с развитием антипиратского законодательства, позволяющего блокировать любые сайты в 2014 году, Российским союзом правообладателей была предложена и другая концепция регулирования гражданско-правового оборота контента, защищенного авторским правом, в сети Интернет – глобальная лицензия.

По задумке авторов законопроекта, каждое электронное устройство конечного пользователя подлежит обложению обязательным сбором за использование и распространение в сети

любых произведений, защищенных авторским правом. При этом на интернет-провайдеров доступа должна быть возложена обязанность собирать причитающееся авторам вознаграждение с самих пользователей и перечислять его на расчетный счет РСП, который будет аккумулировать все денежные средства и в дальнейшем самостоятельно «справедливо» распределять их между авторами всего производимого в мире контента (музыки, видео, текстов, дизайна, ПО, игр).

Несмотря на отторжение предлагаемого механизма “легализации нелицензионного контента в сети” со стороны IT-бизнеса, самих правообладателей и конечных пользователей, концепция была поддержана Министерством культуры РФ, РАО, ВОИС и множеством аффилированных с указанными ОКУПами организаций.

После серьезных протестов интернет-активистов, общественных организаций, правообладателей и Министерства связи удалось добиться отклонения предложенной концепции, получившей в Рунете название “налога на интернет”. В июне 2015 в аппарате Правительства сообщили, что Главное контрольное управление администрации президента сняло с контроля поручение по разработке законопроекта, так как дальнейшая работа над концепцией глобальной лицензии нецелесообразна.

Ситуация вокруг авторского права в сети Интернет позволяет сделать вывод, что в настоящее время существует неразрешимый спор между интернет-сообществом (НКО, пользователи, IT-бизнес), с одной стороны, и корпоративными правообладателями, с другой стороны. Первые обеспокоены слишком жестким регулированием, требуют либеральных реформ и расширения режима общественного достояния, а также способов добросовестного использования. Вторые говорят о том, что старые методы уже неэффективны, и хотят получать доступ к частным коммуникациям граждан, иметь возможность блокировать сайты, привлекать к ответственности нарушителей и возлагать все больше обязанностей и ответственности на информационных посредников.

Проблема несоответствия уровня технологий ограничениям, предусмотренным действующим законодательством об

охране интеллектуальной собственности, а также наличие общественного запроса на более простой доступ к образовательному, развлекательному и научному контенту, привела к формированию альтернативных подходов, основанных на либерализации действующих норм в целях копирования информации, а также предложениям из разных уголков мира по радикальному реформированию законодательства об авторском праве.

Одним из таких предложений стала широкая европейская кампания “Fix Copyright” (“Исправить копирайт”)³³, запущенная четырнадцатью некоммерческими организациями из числа стран-участниц Евросоюза. Авторы кампании обращают внимание на то, что копирайт дисфункционален и не подходит для цифрового мира, его сроки охраны чрезмерно затянуты, а само авторское право в его сегодняшнем виде сбивает с толку пользователей и полностью десинхронизировано с реальностью. Кроме того, правозащитники обращают внимание, что законодательство в этой сфере не гармонизировано и сильно отличается в разных странах Евросоюза.

В 2014 году начался массовый опрос, организованный европейской Пиратской партией и Партией зеленых при Комитете по правовым вопросам Европарламента, связанный с реформой копирайта – “Corywongs”³⁴.

Пользователям со всей Европы предлагалось ответить на ряд важных вопросов, связанных с проблемами применения авторского права в цифровом мире, в том числе по вопросам:

- расширения сферы применения копирайта на использование гиперлинков (ссылок) и агрегаторов ссылок (напр. Google News, Feedly);
- легализации аудио-визуального цитирования как формы добросовестного использования (fair use);
- свободы панорамы (возможность фотографировать и свободно распространять изображения защищенных

33 Кампания по исправлению копирайта // <http://www.fixcopyright.eu/>

34 Кампания по реформе авторского права в Евросоюзе // <http://corywongs.eu/>

копирайтом объектов (зданий, скульптур и т.д.), находящихся в общественных местах);

- отказа от гео-блокинга («балканизация» онлайн пространства, запрет на пользование контентом за пределами страны, в которой он приобретен);
- распространения ответственности за неправомерное использование материалов, защищенных копирайтом, на сервис-провайдеров и веб-платформы.

Это далеко не полный список актуальных проблем, с которыми сталкивается интернет-сообщество, когда речь заходит о применении копирайта в цифровом пространстве. В июле 2015 года 100-страничный отчет, составленный депутатом Европарламента Юлией Реда, с рядом поправок был принят большинством голосов членов комитета. В настоящее время в Европарламенте ведется работа по утверждению предлагаемых поправок.

Необходимо отметить, что, согласно исследованиям ведущих научных учреждений по всему миру, борьба с цифровым пиратством с помощью существующих способов защиты копирайта не способна увеличить продажу контента в сети³⁵. Для увеличения реальных продаж авторского контента принимаемые сегодня меры по ограничению доступа к ресурсам и административному давлению на пользователей и операторов веб-сайтов, а также других информационных посредников неэффективны и даже вредны.

Вместо принятия ограничительных и запретительных правовых мер представляется более правильным создание стимулирующих мер, которые могли бы в действительности повысить уровень цифровых продаж контента как в сегменте B2C, так и в сегменте B2B рынков. Одной из таких мер могло бы стать внедрение системы для идентификации контента и создание технической инфраструктуры, которая могла бы дать широкому кругу пользователей (гражданам и бизнесу) возможность найти

35 Научные исследования в области реформы авторского права и файлообмена // <http://changecopyright.ru/%D0%BB%D0%B8%D1%82%D0%B5%D1%80%D0%B0%D1%82%D1%83%D1%80%D0%B0/>

любой интересующий контент, определить его правообладателя и получить информацию о том, где и на каких условиях можно легально приобрести такой контент.

Возможности реформирования института охраны авторских прав в России

Создание благоприятной среды для инновационного развития, увеличения продаж и гражданско-правового оборота результатов интеллектуальной деятельности

Как видно из предлагаемых в России на высшем уровне концепций охраны авторских прав в сети, участники диалога предлагают прямо противоположные решения для реформирования законодательства о копирайте в цифровую эпоху. Стоит отметить, что в настоящее время в России еще не определена какая-либо единая долгосрочная государственная стратегия в области результатов интеллектуальной деятельности.

Тем не менее, предпосылкой к разработке долгосрочной государственной стратегии в указанной области была и остается Стратегия инновационного развития Российской Федерации на период до 2020 года, утвержденная распоряжением Правительства Российской Федерации от 8 декабря 2011 г. № 2227-р. Вторым документом, на котором базируются все работы в данной области, является Концепция долгосрочного социально-экономического развития Российской Федерации, утвержденная распоряжением Правительства Российской Федерации от 17 ноября 2008 г. № 1662-р.

В целях реализации этой стратегии Минобрнауки предложило принять программу содействия процессам формирования благоприятной для инновационного развития среды, решающую, в частности, следующие задачи:

- нахождение стратегического баланса между защитой интеллектуальных прав и защитой общественных интересов через обеспечение доступа широкой аудитории к материалам, защищенным авторским правом, для обеспечения знаниями, поощрения творчества и роста общественного благосостояния;

- модернизация регулирования использования результатов интеллектуальной деятельности в информационно-телекоммуникационных сетях (в том числе в сети Интернет);
- создание правовых условий для эффективной исследовательской и научной деятельности российских граждан и организаций за счет гибкого определения условий свободного использования результатов интеллектуальной деятельности в исследовательских и научных целях;
- повышение объемов создания и эффективного использования результатов интеллектуальной деятельности, созданных за счет государственных средств, через переориентирование механизмов государственной поддержки творческой активности российских граждан и организаций на наиболее успешные сферы создания и вовлечения в оборот интеллектуальной собственности;
- формирование сбалансированных подходов к ограничению злоупотребления интеллектуальными правами в целях монополизации на рынках интеллектуальной продукции;
- обеспечение полноценного электронного удаленного доступа к фондам российских библиотек и архивов, составляющих национальное достояние всего общества;
- расширение возможностей авторов в распоряжении ими созданными произведениями, в том числе в отношении их права разрешения свободного использования произведений и упрощения процедур их передачи в общественное достояние³⁶.

Изменение модели правовой охраны авторских произведений

Одним из наиболее актуальных вопросов нормативно-правового характера является вопрос о регулировании оборота контента в сети – о выборе правильной правовой модели для предоставления правовой охраны произведениям.

До настоящего времени Бернская конвенция по охране литературных и художественных произведений от 9 сентября 1886 г. (дополненная в последний раз в Париже 24 июля 1971

36 <http://минобрнауки.рф/документы/2881/файл/1405/12.11.30-стратегия-интел.собств.pdf>

г., измененная 2 октября 1979 г.) является основополагающим международным документом в области охраны авторских прав и предусматривает, что правовая охрана произведений предоставляется по умолчанию с момента опубликования произведения. В российском законодательстве выражен еще более свободный подход к предоставлению охраны авторским произведениям. Российское (а до 25 декабря 1991 года – советское) авторское право исходит из того, что правовая охрана предоставляется любым произведениям (как опубликованные, так и неопубликованные) с момента их создания, автоматически, без необходимости выполнения каких-либо формальностей (в частности, без государственной регистрации).

Таким образом, правовая охрана произведения возникает у создателя произведения в момент, как только достигнутый творческий результат облекается в какую-либо объективную форму, обеспечивающую его восприятие другими людьми.

Подобное нормативно-правовое регулирование создает немало сложностей на практике. В связи с отсутствием в российском законодательстве требований о необходимых формальных процедурах, которые позволили бы донести до широкой публики информацию о начале охраны произведения, конечные пользователи и владельцы веб-сайтов, добросовестно полагающие, что авторский контент либо находится в общественном достоянии, либо автор не настаивает на правовой охране исключительных прав на произведение, зачастую становятся фигурантами гражданских и уголовных дел о нарушении исключительных прав.

Указанное правовое регулирование делает невозможным использование «сиротских произведений»³⁷, не позволяет агрегаторам контента и владельцам сайтов определить, какой контент они вправе использовать, а какой нет. Кроме того, отсутствие упорядоченного публичного реестра авторских прав порождает немало сложностей для правоохранительных органов и судов

37 Сиротское произведения – это охраняемые авторским правом работы, для использования которых по закону требуется разрешение правообладателя, но добросовестному пользователю невозможно идентифицировать автора или связаться с ним для получения лицензии.

при возникновении споров о нарушении исключительных прав, когда необходимо оперативно и достоверно определить истинного правообладателя произведения, время его создания и другие существенные характеристики (в том числе объем прав, территорию использования, срок, вид лицензии, по которой было разрешено использование произведений).

К 2015 году стало понятно, что кроме наращивания авторского контента, производимого в промышленном масштабе крупными транснациональными и национальными медиакорпорациями, общее количество контента, производимого человеком, будет увеличиваться в геометрической прогрессии, в связи со значительным развитием сетевых технологий и появлением новых возможностей как для самореализации людей, так и для успешной доставки их произведений до целевой аудитории. Пользователи Livejournal, Youtube, Facebook, Flickr, Instagram и всевозможных иных UGC-сервисов генерируют ежедневно миллионы новых авторских произведений.

Интернет-пространство в цифровую эпоху представляет собой хранилище миллиардов авторских произведений с различным правовым статусом в контексте авторских прав. Отсутствие возможности идентифицировать контент, а также по его идентификатору определить автора контента и иные необходимые сведения для дальнейшего использования контента – все это осложняет работу администраторов веб-сайтов и конечных пользователей. Согласно действующему российскому законодательству, любой гражданин может непреднамеренно стать фигурантом по делу о нарушении исключительных прав на произведение, не зная при этом, кто вообще является действительным владельцем прав на контент и какие права в отношении его использования были истребованы правообладателем.

Кроме того, последние нововведения Федерального закона №364-ФЗ позволяют по требованию правообладателя контента заблокировать навечно доступ к любому веб-ресурсу, что было продемонстрировано в октября-ноябре 2015 года, когда в соответствии с решениями Мосгорсуда было предписано навсегда ограничить доступ к крупнейшими русскоязычным

торрент-индексаторам, включая rutracker.org, rutor.org, seedoff.net, kinozal.tv и др. Анализ формирующейся в настоящее время практики правоприменения показывает, что существующий подход к чрезмерной охране авторских прав с угрозой вечной блокировки сайта и судебного преследования за счет “государственной машины принуждения” порождает иную достаточно существенную проблему современного судопроизводства по авторским спорам – “копирайт-троллинг”. Копирайт-троллингом принято называть недобросовестную деятельность ОКУПов, специализированных юридических агентств и частных лиц, которая выражается в злоупотреблении правом с целью неосновательного обогащения за счет другого лица без надлежащих правовых оснований, под угрозой блокировки сайта, взыскания крупной суммы компенсации или уголовного преследования.

Представляется, что кроме коррекции отдельных положений законодательства об авторском праве для их приведения в соответствие с уровнем современных цифровых технологий и общественными запросами на потребление и копирование информации и контента, назрела необходимость пересмотра фундамента “копирайта” – оснований для предоставления правовой охраны.

Формальные процедуры как основание для предоставления правовой охраны произведению

Парадигма современного гражданского права России в области охраны результатов интеллектуальной деятельности, представляющая авторское право как право интеллектуальной собственности, привела к ложному формированию мнения некоторых защитников жестких законодательных мер, распространяемому ими публично, что указанное право представляет из себя абсолютное право собственности, присущее вещному праву, а, следовательно, порождает право выбора любого способа охраны, даже причиняющего существенный урон общественному благу и базовому праву граждан на беспрепятственный доступ к информации.

Это противоречит самой природе авторского права, как некоего общественного компромисса между частным и публичным интересами. В США при принятии Конституции мысль о том, что

авторы имеют право на монополию копирования, была предложена, но в результате отвергнута отцами-основателями. Вместо этого авторы Основного закона США приняли другую предпосылку – что авторское право является не натуральным правом авторов, а искусственной уступкой, данной им ради прогресса.

Конституция США 1787 г. в ст.1 установила компетенцию Конгресса по обеспечению прогресса науки и полезных искусств путем наделения авторов и изобретателей исключительными правами на их творения в течение определенного срока.

Стоит отметить, что в США, на родине наиболее жестких предложений по реформированию копирайта (в том числе SOPA, PIPA, АСТА, TPP) до 1989 года для предоставления правовой охраны объектам авторского права в США также требовалось соблюдения ряда формальностей.

Законом США 1976 г. устанавливалось обязательное использование отметки об авторском праве, которое уведомляло бы общество о правовой охране произведения, однако это требование было аннулировано, когда Соединенные Штаты стали придерживаться Бернского Соглашения, начавшего свое действие для США с 1 марта 1989 года. Хотя работы, изданные без отметки, до этой даты могли войти в общественное достояние Соединенных Штатов, Uruguay Round Agreements Act (URAA) восстановил авторское право ряда работ, первоначально изданных без отметки.

Использование специальной отметки представлялось важным, так как посредством ее широкая общественность информируется о том, что работа защищена авторским правом, идентифицируется владелец авторского права и отражается год первой публикации. Кроме того, в случае нарушения авторского права на работу, если соответствующая отметка об авторском праве появляется на изданной копии или копиях, к которым имел доступ ответчик в иске по нарушению авторского права, никакого веса не должно придаваться позиции защиты ответчика, основанной на неумышленном нарушении, в смягчении фактического или установленного законом ущерба, за исключением предусмотренного в законе об авторском праве.

Использование отметки об авторском праве являлось ответственностью владельца авторского права и не требовало предварительного разрешения и регистрации в Агентстве по авторским правам.

Предписывалось, что отметка для визуально различимых копий должна содержать три элемента:

1. Символ © (буква С, заключенная в круг), или слово «Copyright» («Авторское право»).

2. Год первой публикации работы. В случае выпуска сборников или производных работ, включающих материалы, изданные ранее, достаточно было указать год первой публикации сборника или производной работы. Год может быть опущен, если живопись, графическая или скульптурная работа сопровождается текстом, если таковой имеется, и воспроизведена на поздравительных открытках, почтовых открытках, почтовой бумаге, драгоценностях, куклах, игрушках или любом другом полезном изделии;

3. Имя владельца авторского права. Имя владельца авторского права, или сокращение, посредством которого распознается имя, или известное альтернативное обозначение владельца.

Например: © 1997 Майкл Джексон

Вместе с тем, закон предписывал, что знак «С», заключенная в круг» используется только на «визуально приметных копиях».

А для фонограмм звуковых записей предусматривалась другая отметка, содержащая три элемента:

1. Символ (буква «Р», заключенная в круг);

2. Год первой публикации звукозаписи;

3. Имя владельца авторского права в звукозаписи, или сокращения, по которым распознается имя, или, обычно, общеизвестное альтернативное обозначение владельца. Если производитель звукозаписи обозначен на ярлыке фонограммы или коробке, и если вместе с отметкой не обозначено какого-либо другого имени, имя производителя должно рассматриваться как часть отметки. Такая отметка об авторском праве должна прикрепляться к копиям или фонограммам таким образом, чтобы «дать доступное

уведомление о притязании на авторские права». Три элемента отметки должны стоять вместе на копиях или фонограммах или на ярлыке фонограммы или коробке.

Подобная маркировка произведений специальным знаком охраны несла двойную функцию:

- предупредительную, так как оповещала потенциальных пользователей этого произведения о том, что данное произведение охраняется в соответствии с национальным законодательством страны;
- информационную, так как конечные пользователи произведения, а также другие заинтересованные лица имели возможность получить информацию о том, кто является обладателем прав на произведения для того, чтобы в случае заинтересованности впоследствии обратиться к автору с предложениями о заключении авторского договора о передаче исключительных либо неисключительных авторских (смежных) прав.

Сегодня любой автор в США может осуществить регистрацию любых опубликованных и неопубликованных произведения в Бюро по авторским правам США (The U.S. Copyright Office) при Библиотеке Конгресса США, однако такая государственная регистрация произведений носит заявительный характер и является лишь правом автора, которым он может воспользоваться в любой момент в течение жизни.

В цифровую эпоху, когда авторских произведений стало настолько много, что нет никакой объективной возможности их все посчитать и каталогизировать, а авторское право в его законотворческом и правоприменительном развитии начало угрожать фундаментальным правам граждан на доступ к информации и распространение контента любым технологическим способом (в т.ч. в файлообменных сетях), представляется правильным изменение критериев для предоставления правовой охраны объектам авторских прав – путем утверждения формальных процедур и прохождения регистрации цифрового объекта/оцифрованного объекта в едином открытом реестре.

Подобная возможность была установлена во Всемирной конвенции об авторском праве, также известной как Женевская конвенция, принятой на Межправительственной конференции по

авторскому праву в Женеве 6 сентября 1952 года (СССР присоединился к ней с 27 июня 1973 г.). Женевская конвенция предусматривает более низкий уровень охраны авторских прав по сравнению с Бернским соглашением. В отличие от Бернской конвенции, исходящей из факта создания произведения как достаточного основания охраны авторских прав на него, Женевская конвенция предусматривает выполнение определенных формальностей для предоставления такой охраны в отдельных странах-участницах. Данные формальности связаны с тем, что в некоторых странах для того, чтобы произведение охранялось, необходимо депонирование экземпляров, регистрация или другие формальные действия.

Однако при принятии Бернской конвенции странами-учредителями Конвенции было установлено, что обладание правами на литературные и художественные произведения, их реализация и судебная защита не требует соблюдения каких-либо формальностей (п.2 ст. 5 Конвенции). Анализ национального законодательства зарубежных стран позволяет сделать вывод, что под “соблюдением формальностей” в области охраны объектов копирайта понимается выполнение определенных процедур, установленных законодательством страны. К их числу можно отнести:

- применение знака охраны авторского права;
- регистрация и депонирование произведений;
- регистрация авторского договора.

Учитывая положения Бернской конвенции, Всемирная организация по интеллектуальной собственности (ВОИС) не связывает охраноспособность произведения с наличием регистрации произведения.

С 1995 года Россия, подписав Конвенцию, стала ее участником и имплементировала в свое национальное законодательство все те же нормы, которые не ставят охрану произведения в зависимость от соблюдения формальностей в виде депонирования объекта или его регистрации. Адаптируя свое законодательство о копирайте с модели “коллективной собственности”, установленной реалиями советского периода, к модели “свободного рынка”, Россия пошла по пути клонирования норм, установленных Бернской конвенцией и рядом западных стран (прежде всего США и Германии).

Так, в п.1 ст. 9 Закона Российской Федерации «Об авторском праве и смежных правах» от 9 июля 1993г. № 5351-1 было прямо предусмотрено, что “для возникновения и осуществления авторского права не требуется регистрации произведения, иного специального оформления произведения или соблюдения каких-либо формальностей”.

Это положение перекочевало и в действующую сегодня 4-ю часть Гражданского кодекса РФ, регулирующего правовую охрану авторских произведений. Согласно п. 4 ст.1269 ГК РФ для возникновения, осуществления и защиты авторских прав не требуется регистрация произведения или соблюдение каких-либо иных формальностей.

Тем не менее, такой порядок существовал в нашей стране не всегда. Согласно п. 1 «Инструкции о порядке проставления знака охраны авторского права на произведениях науки, литературы и искусства, издаваемых в СССР», утвержденной приказом Госкомиздата СССР от 3 июля 1989г. № 212, было установлено, что в соответствии со ст. III Всемирной конвенции об авторском праве, членом которой СССР является с 27 мая 1973г., все охраняемые на территории СССР произведения должны быть обозначены специальным знаком охраны авторского права.

В соответствии с правилами, изложенными в данной инструкции, знак охраны авторского права должен был быть проставлен на всех охраняемых авторским правом произведениях, изданных после 1 августа 1989г., и являющихся первым выпуском в свет данного произведения или перевода. Таким образом, в инструкции содержались императивные нормы, предписывающие правообладателям в обязательном порядке указывать знак охраны авторского права.

Основания для защиты, установленные в Бернской конвенции в области авторских прав, в последний раз пересмотренной в 1971 году в доинтернетовскую эпоху, создают в современном цифровом пространстве массу сложностей для добросовестного использования цифровых произведений. Это также является причиной для увеличивающегося копирайт-троллинга и дискредитации в глазах общества всего института авторского права.

Подход, основанный на защите исключительных прав без необходимости предварительной регистрации, идентификации и депонировании материала, не позволяет ни операторам веб-сайтов, ни конечным пользователям однозначно идентифицировать объект и его правообладателя, а также сделать однозначный вывод о возможности использования произведения без соглашения правообладателя и без выплаты ему вознаграждения.

Так, например, по результатам опроса некоммерческой организации Consumer Focus 73% британских пользователей³⁸ не уверены до конца, что является разрешенным, а что является запрещенным в свете законодательства о копирайте. А согласно опросу Фонда общественного мнения³⁹ 57% российских пользователей, которые готовы платить за контент, не могут отличить легальный контент от нелегального. Подобные цифры вполне обоснованы, т.к. сегодня действительно не существует каких-либо простых способов проверить наличие копирайта на произведение, а также объем прав, который был предоставлен автором или правообладателем для использования принадлежащего ему произведения.

Учитывая современный объем существующего авторского контента в глобальной сети, уровень развития средств коммуникации, сетей и систем передачи данных, создающих принципиально новые возможности для распространения и обмена результатами интеллектуальной деятельности, необходимо инициировать изменения и дополнения важнейших мировых соглашений, регулирующих указанную сферу – Бернскую конвенцию по охране литературных и художественных произведений, Соглашения по торговым аспектам прав интеллектуальной собственности (TRIPS).

Представляется, что законодательство об авторском праве, основанное на принципах, сформированных в XIX-XX веках,

38 Kantar Media (Consumer Focus) To survey British consumers about a wide range of copyright related activities – including those that are legal and illegal under current copyright law. // <http://www.consumerfocus.org.uk/assets/1/files/2010/02/Consumer-Focus-Time-to-change-the-tune1.pdf>

39 Легальный и нелегальный контент в интернете // <http://fom.ru/SMI-i-internet/11164>

успело довольно сильно устареть к 2015 году, когда интернет, предоставляющий широчайшие возможности для копирования любых материалов и быстрой доставки контента, произвел настоящую техническую революцию во многих сферах жизни.

Кроме того, что сформированное в России законодательство в сфере авторских прав не справляется с феноменом неавторизованного копирования в современных социальных и р2р-сетях, оно также создает помехи дальнейшему развитию файлообменных технологий.

В 2014 году такие организации, как “Ассоциация пользователей интернета”, “РосКомСвобода” и “Пиратская партия России” запустили кампанию по реформированию авторского права в России – “Время менять копирайт”⁴⁰. Толчком к этому стало новое антипиратское законодательство, дающее возможность любому правообладателю любого вида контента, защищенного авторским правом по умолчанию, за два и более нарушения своих исключительных прав навечно заблокировать сайт, даже с многомиллионной аудиторией.

Согласно предложениям активистов для балансирования публичных и частных интересов и защиты прав интернет-пользователей предлагается принять ряд законодательных мер стимулирующего характера, которые, в отличие от подхода “антипиратского закона”, могли бы решить актуальные вопросы нормативного регулирования, связанные с противоречием традиционной модели копирайта общественным интересам, в том числе:

- расширение правового режима “общественного достояния” на культурное советское наследие, а также на произведения, создаваемые за счет федерального или муниципального бюджета;
- сокращение сроков охраны авторских прав до разумных пределов;
- декриминализация правонарушений, связанных с неавторизованным некоммерческим копированием и распространением контента в файлообменных сетях;

40 Кампания “Время менять копирайт” // <http://changecopyright.ru/>

- расширение способов добросовестного использования (когда произведение может быть использовано без согласия правообладателя и без выплаты ему вознаграждения);
- правовая регламентация и стимулированием открытых (свободных) лицензий;
- изменение модели коллективного управления авторскими правами;
- ограничение использования DRM;
- регистрация произведений после 5 лет коммерческого использования.

Последний пункт предложения предлагается осуществить путем возврата к положениям Всемирной конвенции об авторском праве (Женевской конвенции 1952 г.), предусматривающей специальный порядок депонирования и регистрации произведения в целях охраны.

Согласно указанному подходу срок охраны исключительных прав на произведение устанавливается на 5 лет с момента публикации. Срок охраны может быть продлен на 5 лет по заявлению правообладателя, поданному в течение последнего года действия этого права. Для дальнейшей пролонгации срока охраны произведения правообладателю необходимо зарегистрировать свои права в публичном реестре авторских прав на произведения по истечении 5 лет с момента публикации. Продление срока действия исключительного права на произведение возможно не более трех раз.

В настоящее время соблюдение формальностей для предоставления правовой охраны авторскому произведению существует в некоторых национальных законодательствах, в т.ч. Испании, Аргентины, Колумбии и др. Условием защиты авторского права на произведение является депонирование экземпляра публикации.

Создание единого реестра объектов авторских прав, а также системы идентификации произведений, в отношении которых истребуется весь объем правовой защиты, предусмотренный копирайтом, представляется крайне важным для всего интернет-сообщества.

Очевидно, что требование о маркировке всех экземпляров произведений специальным знаком охраны в целях информирования общества о защите произведения копирайтом и недопустимости его использования без согласия правообладателя, более не актуально в эпоху, когда все меньшее количество экземпляров публикуется на материальных носителях, а все распространение и использование авторского контента в обществе сводится к цифровому обороту контента в глобальной сети и на цифровых носителях.

Вместо механизма маркировки экземпляров произведений актуальным решением сегодня могла бы стать система уникальной идентификации произведений с присвоением метаданных, а также создание репозитория произведений с релевантными поисковыми возможностями.

При этом, учитывая, что Бернская конвенция не содержит каких-либо положений об ограничении доступа к сайтам в сети Интернет, а также то, что Россия как правопреемник СССР с 1973 года является членом Всемирной конвенции об авторском праве, представляется правильным внесение в российское законодательство об авторском праве изменений, обязывающих правообладателя выполнить ряд формальностей и осуществить регистрацию произведения в случае, если он желает воспользоваться положениями антипиратского закона с возможностью блокировки веб-сайтов за размещение на них нелицензионного контента.

Реестр цифровых объектов как востребованный инструмент фиксации метаданных о произведении и правообладателе

В последнее время в российском законодательстве наблюдается очевидный тренд на введение ограничительных и запретительных законодательных мер в сфере гражданско-правового оборота объектов авторских прав. Большинство из них направлены на поддержание ранее сформированного подхода по защите экономических интересов издателей, так и не сумевших

приспособиться к новым цифровым реалиям. Однако на фоне ужесточения законодательства об авторском праве до сих пор не предпринимается никаких мер стимулирующего характера для реального увеличения цифровых продаж создаваемого в России контента, защиты авторов и увеличении прозрачности контроля за тем, кем и в каком количестве используется контент, и как ОКУПами осуществляется сбор вознаграждения для его дальнейшего перечисления авторам.

Как уже было сказано ранее, в РФ не требуется обязательная государственная регистрация и депонирование произведений для предоставления правовой охраны. Несмотря на то, что регламентация подобных действий отсутствует в действующем российском законодательстве, ничто не мешает правообладателю осуществить такие действия по собственному желанию. Регистрация и депонирование носят заявительный характер и не требуют проведения никакой экспертизы произведения. В настоящее время в России существует множество коммерческих и некоммерческих организаций, оказывающих услуги по депонированию: РАО (Российское авторское общество), Копирус, Республиканский НИИ интеллектуальной собственности и многие другие. Процедура депонирования происходит следующим образом: в регистрирующий орган подаются заявление от автора и само произведение на любом материальном носителе. Произведение прошивается, печатывается либо фиксируется иным способом, далее правообладатель произведения ставит свою подпись, после чего передает экземпляр произведения регистратору. Автору же выдается именное свидетельство на его произведение за определенным номером и с указанием даты выдачи. Таким образом, доказывается то, что произведение одного автора возникло ранее произведения другого автора.

Тем не менее, в последнее время все чаще обсуждается идея создания единого реестра авторских прав на произведения. При этом доводы для создания реестра звучат самые разные.

Вопрос о разработке и принятии стратегии в области результатов интеллектуальной деятельности был впервые поднят Торгово-промышленной палатой Российской Федерации

25 апреля 2012 года на первом заседании Совета по вопросам интеллектуальной собственности при председателе Совета Федерации.

Эту инициативу поддержал Дмитрий Медведев, и к участию в разработке стратегии были привлечены несколько федеральных российских ведомств – Минэкономразвития, Минобрнауки, Минкультуры, Роспатент и МВД.

30 ноября 2012 г. Минобрнауки выставило на общественное обсуждение концепцию основных положений долгосрочной государственной стратегии в области интеллектуальной собственности⁴¹, где, в том числе, говорится о создании открытой системы регистрации объектов интеллектуальных прав и о принятии специальной программы содействия процессам формирования благоприятной для инновационного развития среды, решающей следующие задачи:

- расширение возможностей авторов в распоряжении ими созданными произведениями, в том числе в отношении их права разрешения свободного использования произведений и упрощения процедур их передачи в общественное достояние;
- создание благоприятной среды для творчества, поиска новых научно-технических и конструкторских решений;
- обеспечение доступности информации о результатах интеллектуальной деятельности, облегчение процедур регистрации объектов интеллектуальной собственности;
- способствование формированию организационных и правовых механизмов вовлечения результатов интеллектуальной деятельности в оборот;
- формирование среды эффективных и динамичных собственников и потребителей результатов интеллектуальной деятельности;
- повышение информированности общества в вопросах обращения с интеллектуальной собственностью, в том числе повышение информированности иностранных инвесторов.

41 <http://минобрнауки.рф/документы/2881/файл/1405/12.11.30-стратегия-интел.собств.pdf>

Согласно концепции, для решения задачи открытого представления информации о результатах интеллектуальной деятельности предлагается создать государственные регистрационные системы учета результатов интеллектуальной деятельности, которые должны являться основным информационным источником для участников рынка.

Из указанного документа следует, что для регистрации охраняемых результатов интеллектуальной деятельности предполагается разработать соответствующие сервисы в рамках системы оказания государственных услуг. При этом предполагается поэтапная частичная передача функций по проведению экспертиз аккредитованным негосударственным организациям (что особенно важно с учетом прогнозирования существенного роста количества регистрируемых объектов), а функции контроля соблюдения прав авторов и обеспечения эффективного вовлечения результатов интеллектуальной деятельности в оборот возложить на государство.

В документе также подчеркивается, что необходима разработка механизмов взаимодействия государственных информационных систем по регистрации результатов интеллектуальной деятельности, а также обеспечение межотраслевого взаимодействия с девятью корпоративными системами управления знаниями и нематериальными активами (ГК «Росатом», АО «Роснано», ПАО «Газпром» и иными субъектами хозяйственной деятельности).

Обсуждение этой концепции уже около 3-х лет ведется в Государственной думе РФ и Совете Федерации⁴², однако практическая реализация предлагаемых изменений пока не началась.

В 2013 году Минкомсвязи России предложило⁴³ создать единый реестр авторских и смежных прав (РАСП) – общероссийскую базу данных обо всех фильмах, музыкальных записях, книгах и т.п., а также их правообладателях. Правительство РФ дало поручение Министерству культуры проработать правовую основу

42 <http://минобрнауки.рф/c/новости/5626>

43 Владимир Зыков – “В Рунете все же появится реестр авторских прав”
// <http://izvestia.ru/news/561631>

создания подобного реестра, однако в заключении по рассмотрению вопроса создания РАСП заместитель министр культуры Григорий Ивлиев сделал вывод, что создавать этот реестр нецелесообразно, т.к. для предоставления правовой охраны произведениям не требуется соблюдение никаких формальных процедур. Несмотря на отрицательный отзыв Минкультуры, Минкомсвязи в октябре 2013 г. все-таки опубликовало информацию о конкурсе на создание РАСП. Согласно конкурсной документации, правообладатели должны были получить технологию для создания так называемых цифровых отпечатков для любых принадлежащих им музыкальных или аудиовизуальных произведений, которые предполагалось загружать на официальный веб-сайт РАСП. По замыслу авторов проекта такой реестр позволил бы интернет-площадкам проверять по данным отпечаткам, имеются ли какие-то ограничения на использование аудио- или видео контента, загружаемого самими пользователями. Однако конкурс так и не был проведен и позже был отменен приказом министра⁴⁴.

Согласно п. 5.2 Основных положений долгосрочной государственной стратегии в области интеллектуальной собственности, разработанной Минобрнауки, для решения задачи открытого представления информации о результатах интеллектуальной деятельности должны создаваться государственные регистрационные системы учета результатов интеллектуальной деятельности, являющиеся основным информационным источником для участников рынка. Для регистрации охраняемых результатов интеллектуальной деятельности (получение патента, регистрация программы для ЭВМ и т.д.) предполагается разработать соответствующие сервисы в рамках системы оказания государственных услуг⁴⁵.

Идея разработки реестра была отложена до момента возникновения новой единой федеральной службы по охране

44 Владимир Зыков – «Минкультуры против единого реестра авторских прав» // <http://izvestia.ru/news/563238>

45 Основные положения долгосрочной государственной стратегии в области интеллектуальной собственности // <http://минобрнауки.рф/документы/2881>

интеллектуальных прав. Помимо выработки предложений по функционированию реестра, новому ведомству предписывалось разработать новую стратегию развития интеллектуальной собственности в России, а также решить вопрос о создании реестра научно-исследовательских и опытно-конструкторских разработок (НИОКР), выполненных за счет бюджета.

В июне 2014 г., после принятия Федеральных законов №139-ФЗ и №346-ФЗ («антипиратский закон»), который заметно осложнил доступ граждан к информации, Совет по правам человека при Президенте РФ также внес предложение по созданию реестра авторских прав⁴⁶. По мнению членов Совета, такой реестр необходим, так как он позволит владельцам авторских и смежных прав самостоятельно заносить в него свои произведения. И только после этого можно будет блокировать сайты, где эти фильмы, книги и т.п. незаконно выложены. При этом предполагалось, что реестр будет вести создаваемая Федеральная служба по интеллектуальной собственности.

По предложенной концепции в указанном реестре также необходимо в обязательном порядке зарегистрировать произведения, авторские и смежные права на которые уже перешли в общественное достояние, чтобы исключить возможность их блокировки в интернете.

Кроме того, Совет предложил передать в общественную собственность многие произведения, созданные за государственные деньги. В документе отмечается, что в целях «обеспечения баланса интересов граждан, общества, государства, правообладателей и пользователей необходимо частично передать в общественное достояние культурные ценности, срок охраны которых к 1 января 1993 года истек, а также созданные за счет государства».

В 2015 году вопрос о создании реестра не исчез из повестки дня, но инфраструктура для такой работы так и не была создана.

В конце февраля 2015 года в Общественной палате РФ состоялось заседание экспертной группы по выработке концепции

46 СПЧ предлагает создать список неблокируемых произведений// <http://izvestia.ru/news/572996>

авторского права в цифровой среде, где представители медиа-сообщества предложили ввести мораторий на внесение каких-либо изменений в законодательство об интеллектуальных правах до 2015 года⁴⁷.

В резолюции палаты отмечается, что адресованная Госдуме просьба о моратории связана с вступлением в силу с 1 мая 2015 года поправок, расширяющих действие «антипиратского закона» на другие виды контента, помимо музыки, а также с предлагаемой режиссером Никитой Михалковым «глобальной лицензией».

В документе подчеркивается неправильность выбранного вектора дальнейшего правового регулирования оборота объектов авторских прав, а также отмечается, «что необходимо сформировать правоприменительную практику механизмов блокировки сайтов, а также выработать продуманное решение по стратегии развития законодательства в этой сфере».

В апреле 2015 года после длительного обсуждения на различных площадках идеи регулирования авторского права в цифровой среде через аккредитованную организацию по принципу глобальной лицензии (известной в Рунете как «налог на интернет»), предложенной Российским союзом правообладателей (РСП), в Общественной палате РФ была предложена альтернативная концепция, которая также включала предложение по созданию единого реестра авторских прав.

Инициатива, разработанная членами Общественной палаты при участии «Вымпелкома», «МегаФона» и МТС, подразумевает создание каталога продуктов интеллектуальной собственности, в котором будет храниться информация о правообладателях и их контенте. Чтобы попасть в реестр, правообладателю необходимо будет подать заявку и пройти идентификацию через электронную цифровую подпись. Согласно концепции, «вести каталог может быть поручено Минкульту, Минкомсвязи, Роскомнадзору, компании с 100-процентным государственным участием или аккредитованной организации по управлению

47 Сложные вопросы авторского права // <http://www.oprf.ru/press/news/2015/newsitem/28283>

коллективными правами («наименее желательный вариант»). Информация каталога позволит агрегаторам контента, пользователям и обществам по коллективному управлению правами – в зависимости от итоговой схемы взаимодействия – выходить на авторов или правообладателей и заключать с ними лицензионные договоры по выставленной ими оферте на своих официальных страницах, переход на которые может осуществляться через единый цифровой каталог.

В «дорожной карте» «Медиа и информация. Программа долгосрочного развития интернета»⁴⁸, разработанной Минкомсвязью, Институтом исследования интернета (ИИИ) и Фондом развития интернет-инициатив, в разделе «Модернизация системы управления правами» также содержатся предложения по реформированию авторского права и созданию единого реестра авторских прав.

Так, в плане реализации программы на 1-й квартал 2016 года предполагается, в том числе, ряд конкретных шагов:

- Исключение обязанности выплаты пользователем вознаграждения ОУПДО⁴⁹ в случае наличия прямого договора между пользователем и правообладателем;
- Передача функций по ведению единого каталога (реестра) федеральному органу исполнительной власти (либо альтернативный вариант: допускается существование множества реестров прав, деятельность которых будет контролироваться государством);
- Введение упрощенного порядка судебной защиты авторских и смежных прав и блокировки ресурсов в сети Интернет, если сведения о правообладателе и принадлежащих ему правах занесены в единый каталог (реестр).

Ответственными исполнителями было предложено назначить Государственную Думу РФ, Роскомнадзор, Минкомсвязи и Минкультуры.

48 http://filearchive.cnews.ru/img/cnews/2015/10/23/media_i_informatsiya_dk.pdf

49 Организации по управлению правами на договорной основе (вместо нынешних ОКУПов – Организаций по коллективному управлению правами)

Несмотря на некоторое различие подходов, стоит отметить, что необходимость в подобного рода инструменте для фиксации произведений и связанных метаданных испытывают сегодня все, так как в рамках единой системы в реестр могли бы быть внесены сведения:

- о произведениях, в отношении которых выбран строгий режим копирайта;
- о произведениях, распространяемых по свободным (открытым) лицензиям;
- о произведениях, перешедших в режим общественного достояния.

28 января 2016 года руководитель Федеральной службы по интеллектуальной собственности (Роспатент) Григорий Ивлиев сообщил об окончании работ по созданию национального реестра интеллектуальной собственности (НРИС), главная цель которого – защита зарегистрированных правообладателей. По замыслу создателей реестра, в нем будет содержаться информация обо всех видах результатов интеллектуальной деятельности, в том числе изобретениях и товарных знаках. Для правообладателей вхождение в реестр будет добровольным, и каждый желающий автор сможет внести в него всю необходимую информацию и создать цифровой талон своего произведения. Кроме того, на базе реестра будет создана единая коммуникационная площадка для всех заинтересованных лиц. Потенциальные приобретатели результатов интеллектуальной деятельности при помощи реестра смогут связаться с автором или его представителем и даже провести сделку по приобретению прав на объект⁵⁰.

Реестр как источник информирования о принадлежащих исключительных правах для авторов и корпоративных правообладателей

Несмотря на то, что принято множество мер по усилению правовой охраны произведений, защищенных копирайтом, на сегодняшний день не создано практически никаких механизмов для самого автора, чтобы идентифицировать создаваемые им

50 В России создан национальный реестр интеллектуальной собственности // <http://www.rg.ru/2016/01/28/reestr-site.html>

произведения и информировать о своем авторстве широкий круг заинтересованных лиц. Учитывая современные неограниченные возможности для прямой доставки контента от самого правообладателя до конечного пользователя, представляется разумным создание системы для регистрации произведений авторами (либо уполномоченными ими лицами) и доведения до потребителей и агрегаторов информации об уникальных характеристиках произведения, объеме прав на произведения, условиях и месте приобретения произведения, а также контактных данных автора/правообладателя, с которым может связаться любое заинтересованное лицо с целью заключения договора на использование исключительных прав.

Для корпоративных правообладателей реестр мог бы служить и для других довольно важных функций – управления принадлежащими нематериальными активами и связанности произведений. Так, опыт реализации реестров произведений, таких как CrossRef и EIDR, демонстрирует широкие возможности для продажи и управления контентом в сегменте B2B.

Кроме того, единая система идентификации, учета и поиска контента по уникальным критериям способна повысить для авторов цитируемость и востребованность произведения.

Сам по себе реестр не функционален без репозитория, в которых хранятся сами объекты. Реестр может содержать ссылки на репозитории, где может быть легально размещена оцифрованная копия объекта материального мира или изначально цифровая копия, что позволит хранить и обеспечивать массовый доступ пользователей сети к объектам. При этом сам объект может храниться как в каком-то конкретном репозитории, так и в любом месте в пределах всемирной сети (медиа-магазины, контент-ритейлеры, онлайн-кинотеатры). Правообладатель может указывать все места хранения цифрового объекта, где пользователь может легально и на предлагаемых условиях приобрести интересующий его контент.

Одно и то же произведение может быть представлено в различных цифровых форматах и размерах, с различной скоростью воспроизведения. В таком случае произведение может быть

представлено в виде разных цифровых объектов, метаданные о которых будут также различаться в реестре.

Реестр как источник информирования о метаданных для правоприменителя

Координационный совет по защите интеллектуальной собственности Минкомсвязи при обсуждении реестра авторских прав отметил, что появление и активное использование такого реестра является значимым для совершенствования дальнейшей работы правоохранительных органов и исполнительной власти по разработке конкретных мер воздействия на нелегальное распространение произведений в интернете.

Учитывая, что авторское право охраняется с момента создания и не требует никаких формальных процедур для охраны, в правоприменительной практике все чаще возникают проблемы, связанные с подтверждением авторства и идентификацией правообладателя, как надлежащей стороны по делу. Согласно действующему законодательству за защитой нарушенных прав в суды и уполномоченные органы исполнительной власти могут обращаться как сам первоначальный правообладатель (лицензиар) и его представители (некоммерческие организации, представляющие иностранных правообладателей, например РАПО, НФМИ и др.), так и производные правообладатели (лицензиаты), обладающие правами на территории РФ на основании лицензионных договоров.

При этом необходимо учитывать, что лицензионный договор может быть заключен в отношении передачи только отдельных видов прав, а также на ограниченную территорию и срок. Зачастую права передаются хозяйствующим субъектам на территории РФ посредством цепочки сублицензионных договоров, в которой могут быть задействованы и организации с оффшорными юрисдикциями (как правило, это делается в целях налоговой оптимизации). Кроме того, могут передаваться не все права на сложный объект (например, компиляция или аудиовизуальное произведение), а лишь отдельные права на составной объект.

Все это создает сложности в проверке правомочности притязаний на охрану исключительных прав и правоспособности правообладателей, участвующих в сублицензионных цепочках.

В настоящее время не существует никаких действенных способов проверки заявленных прав, кроме изучения представленных договоров и учредительных документов лиц, заявляющих свои притязания на обладание теми или иными произведениями. Правоприменитель также не имеет никакой реальной возможности получить и проверить первичную информацию о произведениях, в отношении которых правообладатель истребует административную и судебную защиту, кроме рассмотрения сведений, которые предоставляет сам заявитель.

Реестр как источник информирования о правообладателях для информационных посредников (сервис-провайдеров) и конечных пользователей

В декабре 2012 г. на заседании Координационного совета по защите интеллектуальной собственности Минкомсвязи РФ представило концепцию⁵¹ реестра авторских прав и обсудило ее с самими правообладателями.

В связи с тем, что сами производители контента призывали законодательно возложить ответственность за нелегальное распространение информации не только на владельцев сайтов, но и на провайдеров, на защиту последних встало Минкомсвязи РФ, выразив свою позицию в письме на имя председателя Координационного совета по защите интеллектуальной собственности Е.В. Савостьянова.

В письме справедливо отмечается, что условие о возможной осведомленности провайдеров о содержании передаваемого ими материала вступает в противоречие с действующими нормами законодательства. Минкомсвязи полагает, что провайдер не может знать о содержании передаваемых материалов, будучи лишь посредником в передаче информации.

Использование реестра данных о произведениях, имеющих ограничения к свободному распространению и использованию в сети Интернет, могло бы стать решением проблемы. В этом случае распространители информации во избежание дальнейших претензий со стороны правообладателей и их представителей, а также в целях

51 Представители правообладателей одобрили концепцию реестра прав от Минкомсвязи // <http://www.copyright.ru/ru/news/main/2012/12/24/registry/>

минимизации рисков привлечения к ответственности имели бы возможность обращаться к реестру для уточнения сведений, подтверждающих или опровергающих статус свободного или ограниченного доступа к защищенному авторским правом контенту.

Для этого в реестр необходимо вносить основные данные о произведении и его правообладателе. Согласно предложению Министерства, «включение данных в реестр для правообладателей должно быть осуществлено сугубо на добровольных началах. Любой сайт и провайдер, желающие использовать зарегистрированный в реестре контент, смогут обращаться к нему для уточнения данных о произведении, а добавив информацию о своих правах на контент в реестр, правообладатель обезопасит себя от открытого пиратства».

Представители операторов связи и Ассоциация «Интернет-Видео», представляющая дистрибьюторов цифрового видеоконтента, поддержали⁵² указанную концепцию, указав, что это может значительно снизить существующие издержки. Интернет-отрасль, НКО, Минкомсвязи и Общественная палата также позитивно отреагировали на представленную концепцию.

Вместе с тем, заместитель главы Минкомсвязи Алексей Волин выразил свое мнение⁵³ о том, что подобных реестров должно быть несколько – чтобы на рынке была конкуренция. По словам Волина, ничто не мешает создать подобные реестры. При этом он убежден, что помощь государства для создания реестра отрасли не нужна. Если появится работающая модель, можно будет думать, как придать ей государственный статус.

Принципы функционирования реестра как системы идентификации контента: унификация, открытость, возмездность.

Идентификация, учет и каталогизация авторских произведений является важным шагом для всех сторон, участвующих

52 Каталог попал под блок // <http://www.kommersant.ru/doc/2713411>

53 Общественная палата предложила альтернативу «налогу на интернет» // <http://izvestia.ru/news/588280>

в процессе оборота и потребления контента. Как было указано ранее, присвоение контенту, обладающему творческой и научной ценностью, уникального идентификатора, позволяющего однозначно определить само произведение, его автора и иные важные характеристики, присущие произведению определенного рода, является необходимой задачей, которая должна быть реализована для самих авторов, операторов веб-сайтов, органов государственной власти и конечных пользователей.

Создаваемый реестр, который будет служить инструментом для получения информации о любом произведении, должен отвечать ряду основополагающих принципов: *унификации, открытости, возмездности.*

1. Прежде всего работа реестра должна строиться *на принципе унификации*, то есть приведения информации к единообразной форме. Наличие однотипного программного обеспечения, единого алгоритма идентификации контента и его однозначного местонахождения в сети Интернет способно обеспечить эффективное взаимодействие между различными системами без дополнительных затрат на совместимость систем и баз данных.

2. *Открытость (или публичность)* реестра означает его доступность широкому кругу лиц, заинтересованному в получении доступа к информации о произведении. Реестр должен позволить любому заинтересованному лицу получить доступ к информации о контенте по его уникальному идентификатору либо по метаданным, в том числе названию произведения, автору произведения, а также иным критериям, необходимым для поиска релевантной информации.

3. Несомненно, использование указанной системы возможно лишь на бесплатной основе для широкого круга лиц, однако, очевидно, что инфраструктура реестра, а также деятельность регистраторов произведений должна быть материально обеспечена. Таким образом, представляется справедливым *взимание платы* за регистрацию и присвоение идентификатора произведению с самих правообладателей либо уполномоченных ими лиц (например, издательств, студий звукозаписи, продюсерских компаний). При этом регистрации должен единожды подлежать сам

правообладатель контента, а также все публикуемые и вносимые им в реестр однозначно идентифицируемые произведения.

Анализ мировой практики по созданию и функционированию цифровых реестров (на примере архитектуры DOI)

Появление DOI как системы идентификации контента

Как видно, предложения по созданию цифрового реестра авторских и смежных прав звучат в последнее время с самых разных сторон, однако до сих пор ни одним из участников взаимоотношений по обороту цифровых произведений не сформулировано техническое решение для создания подобного реестра, не определены стандарты и принципы его работы, а также порядок взаимодействия субъектов с таким реестром.

Вместе с тем технические решения для цифровой идентификации объектов уже существуют в мировой практике. Одним из наиболее распространенных и наиболее востребованных инструментов для идентификации произведений в сети Интернет является *DOI – Digital Object Identifier*, Цифровой Идентификатор Объекта, представляющий из себя признанный международный стандарт представления информации об объекте.

Предпосылкой для создания DOI стало понимание того, что система адресации машин (IP-адресация), являющаяся основой глобальной сети, после бурного развития сети и увеличения в геометрической прогрессии контента перестала соответствовать требованиям сообщества по отслеживанию движения информации. По этой причине доктор Роберт Кан (являющийся одним из изобретателей протокола TCP, которого заслуженно называют одним из отцов интернета) предложил перестроить способ обработки информации, передаваемой по сетям.

Как неоднократно указывал автор системы DOI г-н Кан, интернет – это не кабели, маршрутизаторы и коммутаторы, которые приносят контент на компьютер конечного пользователя. Интернет – это «архитектура» или дорожная карта для подключения в единой сети компьютеров друг к другу, это система

разбивки информации на пакеты, которые передаются путем перенаправления их с сервера на сервер по всему миру, пока они не достигнут места назначения, где они повторно соберутся в определяемый пользователем контент.

Понятие «коммутация пакетов» (packet switching) было разработано в 1960-х годах и стало реальностью в начале 1970-х годов, когда правительственной организацией США, известной как DARPA, впервые была запущена сеть ARPANET, подключившая 23 университета и правительственные компьютеры друг к другу. Когда число компьютеров к середине 1970-х годов значительно выросло, стало ясно, что машины в сети должны иметь общий язык для передачи информации посредством пакетов. Доктор Кан, который в тот момент работал в DARPA, и Винт Серф, второй «отец интернета», работавший вместе с ним, придумали язык, называемый TCP/IP, что расшифровывается, как протокол управления передачей (TCP) и протокол интернета (IP).

Тот интернет, который мы знаем сегодня, появился 1 января 1983 года, когда ARPANET начала использовать TCP/IP. Тем не менее, TCP/IP-адресация была создана, чтобы указать на то, «куда отправлять», а не «что отправлять». Протоколы передачи данных умышленно задумывались «слепыми» к содержанию пакетов информации. На тот момент основной задачей было решение вопроса, как передать информацию из одного места в другое и как соединить между собой многочисленные локальные сети.

Для адресации веб-сайтов в сети Интернет используется IP-адрес. Но IP-адрес сможет привести лишь на домашнюю страницу сервера, с помощью него нельзя найти именно то, что на данном сайте интересует пользователя. Систему адресации к отдельным страницам сайтов в сети, которую мы используем ежедневно для поиска конкретной информации, представляют «универсальные указатели ресурсов» или URL, состоящие из двух компонентов: имени хоста и пути к файлу.

Несомненно, протокол TCP/IP – это прекрасный способ передать информацию между компьютерами, однако он совершенно не пригоден для управления контентом на этих компьютерах.

Основная проблема заключается в поиске информации. Даже если известен точный адрес, находящийся в закладках браузера пользователя, нет никакой гарантии, что он будет работать в следующем году или даже на следующей неделе: владелец веб-ресурса может просто перенести контент на другой ресурс или на другую страницу. Учитывая то, что существующая система адресации не предназначена для распознавания содержания передаваемых пакетов информации, не существует какого-либо простого способа идентификации и защиты данных, которые охраняются копирайтом.

Принимая во внимание все вышесказанное, г-н Кан пришел к выводу о необходимости разработки новой системы для работы с контентом в сети Интернет, которая будет выступать в качестве «верхнего слоя» существующей инфраструктуры, при этом будет решать вопрос не адресации передаваемого пакета, а содержания информации в нем.

Результатом его работ стала так называемая «система указателей» («handle system»). Вместо выявления места отправления или назначения файла с информацией, система присваивает указатель/идентификатор («handle») самой информации, так называемому «цифровому объекту». Цифровой объект – это все, что может храниться на компьютере: отдельная веб-страница, документ, музыкальный файл, видеофайл, книга или глава книги или даже стоматологический рентген. Главной особенностью этой системы, связывающей ее с интернетом, является наличие метаданных, «привязанных» к каждому идентификатору. Именно в метаданных хранятся сведения, относящиеся к объекту, в том числе и адрес в сети интернет (URL), где можно найти цифровой объект. За этой универсальной технологией закрепилось название Digital Object Architecture (DOA).

Так же, как имя хоста решает, к какому IP-адресу обратиться, handle решает вопрос, какую информации об объекте должен знать компьютер. Handle позволяет “объяснить” компьютеру, какой именно объект содержится в файле, как часто он будет обновляться и как именно разрешается использование указанного объекта. Как и в системе DNS, цифровой адрес уполномоченного

субъекта может быть представлен в текстовом виде, что обеспечивает более простое запоминание.

В 1994 году Ассоциация американских издателей почувствовала революцию, которую совершат электронные книги, и решила, что нужен какой-нибудь способ идентифицировать контент и настроить некоторые базовые средства для защиты принадлежащих Ассоциации авторских прав. В 1997 году Ассоциация начала работать с профессором Каном и CNRI для того, чтобы адаптировать систему “handle” и DOA в целом под свои нужды. Результатом этого стал каталог DOI.

DOI возникла в результате совместной инициативы трех торговых организаций книгоиздательской индустрии: Международной ассоциации издателей, Международной ассоциации научных, технических и медицинских издателей и Ассоциации американских издателей.

Впервые DOI была представлена на книжной выставке во Франкфурте в 1997 году. В этом же году был основан Международный фонд DOI (IDF) для развития и управления системой. С самого начала процесса IDF работал вместе с Корпорацией национальных научно-исследовательских инициатив, которая выступала в качестве технического партнера, применяя систему дескрипторов (идентификаторов системных объектов) Handle System (где «хендл» – это номер адреса в системной таблице некоторой структуры данных), разработанную в CNRI.

CNRI так и остается партнером фонда IDF. С 1998 года Фонд тесно сотрудничает с проектом *indexs* (совместимость данных в системе электронной коммерции; 1998-2000) и рядом других инициатив.

Описанный выше стандарт начал использоваться для идентификации и связности печатных произведений – от отдельной журнальной статьи до книги. Он мог включать и другие международные коды идентификаторов: ISSN, ISBN, ISTC. Первое применение системы DOI, связывание ссылками электронных статей (объектов) органом регистрации перекрестного обмена, было осуществлено в 2000 году.

Несмотря на то, что система берет свое начало в книгоиздательстве, со временем DOI стала использоваться не только для регистрации текстовых авторских произведений, но и других видов произведений, представленных в сети: аудиозаписи, аудиовизуальные произведения, фотографии, таблицы, препринты, рисунки, а также иные произведений, представленные в цифровом либо аналоговом виде.

В 2010 году система DOI была признана соответствующей стандартам ISO, в результате чего системе был присвоен стандарт ISO 26324, Digital Object Identifier System (Система цифровых идентификаторов объектов). Данный международный стандарт определяет синтаксис, описание, функциональные компоненты разрешения системы цифровых идентификаторов объектов, а также общие принципы для создания, регистрации и управления имен в DOI.

Вместе с тем он не выделяет отдельные технологии для реализации синтаксиса, описания и функциональных компонентов разрешения системы цифровых идентификаторов объектов. ISO 26324 является инструментом, при помощи которого система DOI была утверждена в качестве международного стандарта, а IDF получил название Регистрирующего органа (администрация) ISO 26324 Registration Authority.

Синтаксис DOI является стандартом Национальной организации по информационным стандартам (US), ANSI/NISO Z39.84-2010. Впервые он был опубликован в 2000 году, пересмотрен и исправлен в 2005 году с целью устранения чувствительности к регистру, что не было реализовано ранее, таким образом, при внесении данных изменений не возникало проблем совместимости с предыдущими версиями. Современный стандарт ISO 26324 (2012) включает описание синтаксиса. Стандарт ISO также охватывает расширяемую схему метаданных и гарантии в области сохранности данных и т.д., которые не входили в стандарт ISO.

DOI и более широко DOA могут выступать в качестве унифицированных указателей ресурса в пределах пространства имен info-URI (IETF RFC 4452, информационная структура для

информационных ресурсов с идентификаторами в публичных пространствах имен).

Сегодня DOA/DOI представляет собой уникальную систему, обеспечивающую инфраструктуру для постоянной идентификации объектов любого типа. DOA/DOI может определить постоянное местонахождение произведения в сетевом пространстве, его наименование, имя автора, а также иные метаданные, описывающие произведение.

Имя DOA/DOI является постоянно присвоенным объекту с целью предоставления разрешаемой постоянной ссылки на текущую информацию об указанном объекте. В то время как информация об объекте может со временем изменяться, его имя DOI остается неизменным. В пределах системы имя DOI может быть разрешено в значения одной или более разновидностей данных, относящихся к объекту, идентифицированному указанным именем DOI, таким как URL, адрес электронной почты или другими идентификаторами и описательными метаданными.

Система DOA/DOI разрешает формирование автоматизированных систем и транзакций. Режимы использования системы DOI включают в себя: управление информацией, размещение документации и доступа, управление метаданными, упрощение электронных транзакций, постоянную уникальную идентификацию любых форм информации для использования в коммерческих и некоммерческих операциях.

Содержание (контент) объекта, связанное с именем DOA/DOI, четко описывается метаданными согласно структурной расширяемой модели данных, что позволяет объекту быть связанным с метаданными любой необходимой степени точности и детализации. Модель данных позволяет установить операционную совместимость и взаимодействие между отдельными приложениями.

Границы системы DOA/DOI определены ссылкой не на тип (формат и пр.) контента объекта, а ссылкой на функциональные возможности, которые он предоставляет, а также на контекст его использования.

Организационная структура IDF

Фонд IDF – некоммерческая компания, зарегистрированная в штате Делавэр, США. Операционные издержки фонда компенсируются за счет системы членских взносов. Фонд также уполномочен проводить регистрацию согласно стандарту ISO 26324, в соответствии с которым IDF координирует и управляет системой DOI, а также предоставляет инфраструктуру для ее функционирования, регистрирует префиксы имен и т.д. Непосредственной регистрацией и поддержкой объектов занимается объединение регистрационных агентств, которые следуют общей политике IDF. Кроме агентств, в IDF Foundation существуют еще три типа членства: общее, привилегированное и аффилиат. Все категории, кроме последней, обладают правом голоса внутри фонда. Общее членство предоставляется организациям, поддерживающим разработку системы DOI, но не являющимся регистрационным агентством. Привилегированное членство резервируется для основателей IDF и организаций правообладателей. Аффилированное членство ограничено профессиональными ассоциациями, в составе которых есть члены IDF. Все привилегированные члены представлены в совете фонда, регистрационные агентства – через год после получения статуса, общие члены – одним выборным представителем на всех.

Наиболее важным участником в системе являются регистрационные агентства, непосредственно оказывающие услуги распределения префиксов, регистрации имен и предоставления инфраструктуры для объявления и поддержания метаданных объектов. Им присваивается статус Registration Agencies, «Агентств Регистрации»⁵⁴ (далее – регистраторы). При этом регистрацией имен в системе DOI также может заниматься любая организация или физическое лицо по договору с регистрационным агентством. Такая организация не обязана быть членом IDF.

54 DOI Registration Agencies // http://www.doi.org/registration_agencies.html

Основные задачи регистраторов: услуги по выделению префиксов DOI, регистрация имен DOI и обеспечение необходимой инфраструктуры, позволяющей поддерживать функционирование системы регистрации объектов, запись данных и метаданных. IDF следит, чтобы регистраторы поддерживали целостность системы DOI на максимально качественном уровне и осуществляли стабильную и надежную выдачу данных для пользователей DOI. Данные, предоставляемые пользователям от регистратора, должны быть точны, а метаданные согласованы и соответствуют как ядру системы DOI, так и соответствующим принятым стандартам прикладного уровня.

С момента создания DOI различные сферы (например, управление активами индустрии развлечений, сфера публицистики, сфера контента на национальных языках) стали контролироваться различными регистраторами.

На данный момент IDF указывает о наличии 10 официальных регистраторов DOI:

Наименование	Адрес сайта	Представительство в странах и регионах
Airiti, Inc.	doi.airiti.com	Тайвань
CrossRef	crossref.org	Великобритания и США
China National Knowledge Infrastructure (CNKI)	eng.oversea.cnki.net	Китай, Тайвань, Япония, Южная Корея, Северная Америка, Азиатско-Тихоокеанский регион
DataCite	datacite.org	Великобритания, Австралия, Канада, Китай, Дания, Эстония, Франция, Германия, Венгрия, Италия, Япония, Нидерланды, Норвегия, ЮАР, Южная Корея, Швеция, Швейцария, Таиланд, США
EIDR (Entertainment Identifier Registry)	eidr.org	США

ISTIC (The Institute of Scientific and Technical Information of China)	doi.org.cn	Китай
JaLC (Japan Link Center)	japanlinkcenter.org	Япония
mEDRA (Multilingual European DOI Registration Agency)	medra.org	Италия
OP (Publications Office of the European Union)	publications.europa.eu	Европейский Союз
KISTI (Korea Institute of Science and Technology Information)	http://www.kisti.re.kr	Корея

Каждый регистратор сам устанавливает политику взаимодействия с издателями, библиотеками и пользователями, которые в той или иной степени используют систему DOI.

Различие подходов к регистрации и предоставления доступа к системе DOI проявляется в следующем:

- в ценовой политике доступа к системе регистрации объектов в системе DOI;
- в языковой области;
- в области знаний;
- в создании дополнительных инструментов, наращивающих возможности DOI⁵⁵.

Сам Фонд IDF так определяет области интересов официальных регистраторов DOI⁵⁶:

- Airiti, Inc.

55 Registration Agencies – Areas of Coverage // http://www.doi.org/RA_Coverage.html

56 DOI System Examples // <http://www.doi.org/demos.html>

Сфокусирован на материалах, опубликованных на традиционном китайском языке;

- CrossRef⁵⁷.

Ориентирован на контент, связанный с научными и профессиональными исследованиями, а также на статьи, книги, материалы различных конференций. В приоритет ставится удобный поиск по базам метаданных материалов, опубликованных в цифровой форме;

Примеры.

DOI-имена в контексте ссылок на научно-техническую журнальную статью и главу книги:

- Journal Article: «Quantum tomography: Measured measurement», Markus Aspelmeyer, nature physics January 2009, Volume 5, No 1, pp11-12;

[doi:10.1038/nphys1170]

- Book Chapter: Held, Gilbert. Internetworking LANs and WANs (Second Edition), John Wiley & Sons, 1998, Published Online 05 Oct 2001. Chapter 1, Network Concepts (p 1-30)

[doi:10.1002/0470841559.ch1]

- China National Knowledge Infrastructure (CNKI)

Направлен на китайско-ориентированные информационные ресурсы, в том числе в области политики, экономики, гуманитарных, социальных и иных наук и технологий. CNKI публикует базы, содержащие электронные версии журналов, газет, диссертаций, научных трудов, ежегодников, рефератов и др.;

- DataCite

Ориентирован на улучшение научной инфраструктуры вокруг наборов различных данных. Проводит мероприятия по обмену передовой практикой, выявлению и решению ряда уникальных проблем, которые возникают при работе с наборами данных. DataCite работает с дата-центрами и организациями, которые аккумулируют данные. Регистратор использует бизнес-модель,

57 Загрузка данных в CrossRef // http://elpub.ru/images/files/Depositing_to_CrossRef.pdf

которая отвечает потребностям как небольших коммерческих организаций, так и крупных национальных;

Примеры.

DOI-имена для доступа к зарегистрированным научным наборам данных:

Множества и подмножества:

- Irino, T; Tada, R (2009): Chemical and mineral compositions of sediments from ODP Site 127-797. Geological Institute, University of Tokyo.

[doi:10.1594/PANGAEA.726855]

Сообщение о землетрясении от автоматизированной системы:

- Geofon operator (2009): GEOFON event gfz2009kciu (NW Balkan Region) GeoForschungsZentrum Potsdam (GFZ).

[doi:10.1594/GFZ.GEOFON.gfz2009kciu]

Визуализация набора данных на карте:

- Kraus, Stefan; del Valle, Rodolfo (2008): Geological map of Potter Peninsula (King George Island, South Shetland Islands, Antarctic Peninsula). Instituto Antártico Chileno, Punta Arenas, Chile & Instituto Antártico Argentino, Buenos Aires, Argentina. [doi:10.1594/PANGAEA.667386]

Видео проведения глазной операции как дополнительный материал к научному журналу:

- B. Kirshhof (2009) Silicone oil bubbles entrapped in the vitreous base during silicone oil removal, Video Journal of Vitreoretinal Surgery.

[doi: 10.3207/2959859860]

- EIDR (Entertainment Identifier Registry)

Ведет реестр фильмов, телевизионных шоу и других коммерческих аудио- и видеоактивов. Регистратор создан для вещательных и кабельных сетей, предоставляет идентификаторы и связанные метаданные, которые применяются в рекламной киноиндустрии, от этапа редактирования отснятого материала до широковещания, цифровой дистрибуции и составления отчета;

- ISTIC (The Institute of Scientific and Technical Information of China)

Предоставляет услуги для китайских журналов, диссертаций, книг, материалов конференций и других литературных ресурсов, а также осуществляет менеджмент подборок научных данных, мультимедийных аудио- и видеоресурсов в Китае;

Примеры.

DOI-имена китайских журналов, наборов данных и диссертаций:

- Журнальная статья:

张愚. 氢键复合物中键长变化与振动频率移动相关性重访 [J]. 物理化学学报 2012, 28(03):499-503. [doi:10.3866/PKU.WHXB201112303]

- Научные данные:

阎广建, 康国婷, 任华忠等. 黑河综合遥感联合试验: 盈科绿洲、花寨子荒漠和临泽草地加密观测区地基热像仪地表辐射温度观测数据集. 北京师范大学; 中国科学院遥感应用研究所; 中国科学院地理科学与资源研究所; 兰州交通大学. 2008. [doi:10.3972/water973.0145.db]

• Yan Guangjian, Kang Guoting, Ren Huazhong, Chen Ling, He Tao, Wang Haoxing, Wang Tianxing, Liu Qiang, Li Hua, Xia Chuanfu, Zhou Chunyan, Chen Shaohui, Yang Tianfu. WATER: Dataset of LST (land surface temperature) observed by the thermal camera in the Yingke oasis, Huazhaizi desert steppe and Linze grassland foci experimental areas. Beijing Normal University; Institute of Remote Sensing Applications, Chinese Academy of Sciences; Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences; Lanzhou Jiaotong University. 2008. [doi:10.3972/water973.0145.db]

• Диссертация: 刘乃安. 生物质材料热解失重动力学及其分析方法研究 [D]. 安徽: 中国科学技术大学, 2000. [doi:10.7666/d.y351065]

- JaLC (Japan Link Center)

Оказывает общественные информационные услуги для развития науки и технологий в Японии. Агрегирует научно-учебные

метаданные и контент со стороны национальных институтов и университетов, таких как Japan Science and Technology Agency, National Institute of Informatics, National Diet Library and National Institute of Material Science и других государственных и коммерческих организаций и издателей. JaLC сотрудничает с CrossRef для регистрации в DOI англоязычных журналов в Японии;

Примеры.

DOI-имена японских журнальных статей:

- 竹本 賢太郎, 川東 正美, 久保 信行, 左近 多喜男, 大学におけるWebメールとターミナルサービスの研究, 標準化研究 Vol.7(2009), No.1 p.11-20

[doi:10.11467/iss2003.7.1_11]

- 川崎 努, 植物における免疫誘導と病原微生物の感染戦略, ライフサイエンス 領域融合レビュー, 2, e008 (2013), [doi:10.7875/leading.author.2.e008]

- mEDRA (Multilingual European DOI Registration Agency)

Организует систему цитат для интернет-документов. Отслеживает связь между объектами интеллектуальной собственности. Выдает сертификаты цифровой подписи;

Примеры.

DOI-имена с использованием mEDRA:

- Журнальная статья:

Prodi, Romano. «L'industria dopo l'euro», L'industria-Rivista di economia e politica industriale 4, 559-566 (2002); [doi:10.1430/8105]

- Монография:

Attanasio, Piero. «The use of DOI system in eContent value chain: The case of Casalini Digital Division and mEDRA», White Paper (PDF). [doi:10.1392/BC1.0]

- OP (Publications Office of the European Union) является официальным издателем институтов, органов и учреждений Европейского Союза. Данный регистратор несет ответственность за присвоение DOI от имени этих организаций в ЕС,

включая идентификацию всех монографий ЕС, официального журнала ЕС и его отдельных актов, а также ряда научных статей.

Следует отметить, что по состоянию на июнь 2014 года подавляющую долю рынка, порядка 72% регистраций объектов в системе DOI, занимал только один регистратор – CrossRef. Приложение CrossRef используют более 4800 издателей и обществ в целях обеспечения перекрестного цитирования для научных публикаций. Еще один, ISTIC (The Institute of Scientific and Technical Information of China), охватывал собой порядка 21%. На остальных регистраторов в совокупности приходилось менее 7% от всей деятельности по регистрации объектов в системе DOI⁵⁸:

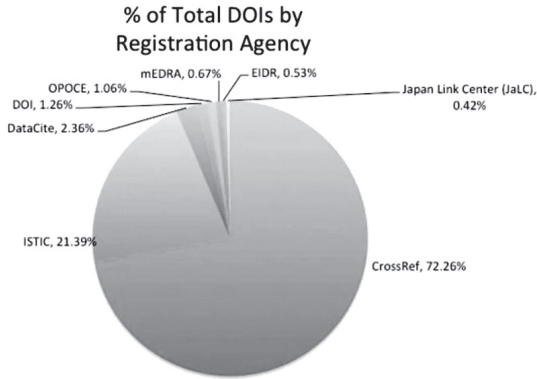


Рисунок 6. Доли RA на рынке

Пользователи могут участвовать в предоставлении услуги, предлагаемой регистрирующим агентством DOI, посредством регистрации материала в одном из агентств или формированием сообщества для создания такой услуги. Использование существующих имен DOI может быть бесплатным. Стоимость регистрации новых имен DOI зависит от услуг, использующих DOI, которые предоставляются регистрирующим агентством. Любое регистрирующее агентство может предлагать свою собственную модель бизнес-процесса в соответствии с общей политикой

58 CrossRef DOIs for Books Webinar – June 26, 2014 // <https://youtu.be/vZmZoPMfoIA>

DOI. Отдельные регистрирующие агентства принимают соответствующие правила для своих сообществ и приложения.

Принимая во внимание то, что система DOI создана для обеспечения осведомленности сети и взаимодействия (операционной совместимости) между участниками, не представляется сложным сформировать любой комплекс современных приложений при помощи DOI-архитектуры. К примеру, система DOI применяется во внутренних процессах многих компаний для издательской деятельности и составления отчетности, невзирая на корпоративные и государственные границы, а также в сфере семантических веб-приложений, находящейся в настоящее время на стадии становления.

Управлением всей инфраструктурой DOA занимается некоммерческая организация The DONA Foundation, зарегистрированная в 2014 году в Женеве, Швейцария. Руководство DONA придерживается мультистейкхолдерского подхода в управлении; согласно швейцарским законам, фонд является автономной организацией и не может быть под контролем правительственных структур.

Целью DONA является управление, разработка, технологическая координация и использование в общественных интересах архитектуры DOA по всему миру, включая ее логические расширения. Для этих целей DONA занимается администрированием регистрации и резолюции идентификаторов цифровых сущностей, то есть управлением распределенным Глобальным реестром хэндлов (GHR), ключевым компонентом архитектуры. В эту задачу входит также аккредитация и координация GHR между несколькими равноправными первичными администраторами (Multi-Primary Administrators, MPAs). Первым таким администратором стала CNRI, ранее управлявшая GHR единолично, двумя другими – китайская Coalition for Handle Services (ETIRI, CDI, CHC) и немецкий Gesellschaft für Wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG). Четвертым MPA де-факто является Международный Союз Электросвязи, которому отведена особая роль фактора стабильности системы. Пятым регистратором должен стать International DOI Foundation,

объединив, таким образом, в общей организационной структуре DOA и DOI. Два представителя МРА (но не более трети от общего количества) входят в правление DONA, другие члены правления выбираются согласно их вкладу в архитектуру DOA и экспертизе в соответствующих технологических областях.

Организации, желающие предоставлять услуги идентификации и резолюции в рамках DOA, регистрируются у МРА в качестве Local handle service provider, получают свой префикс в системе, устанавливают и настраивают соответствующее ПО. В рамках DOA все регистрационные агентства системы DOI также являются провайдерами LHR.

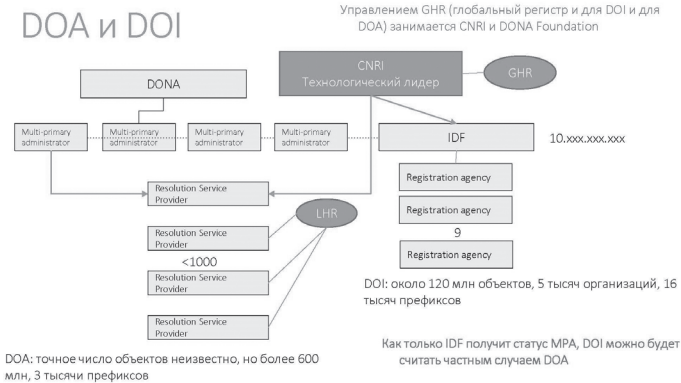


Рисунок 7. Организационная структура DOA

Примеры и перспективы использования идентификации цифровых объектов

Использование DOI значительно облегчает цитирование, поиск и локализацию публикаций и произведений. В настоящее время DOI является неотъемлемым атрибутом системы научной коммуникации за счет эффективного обеспечения процессов обмена научной информацией. Опыт мирового научного сообщества представляется очень ценным и позволяет рекомендовать использование DOI и во всех других творческих отраслях.

По состоянию на декабрь 2015 года в системе DOI регистрационным агентствам выдано более 15 тысяч префиксов, другими МРА DOA – еще около 3 тысяч. Количество зарегистрированных объектов DOI – более 114 млн. Точное количество объектов

DOA неизвестно, так как несколько миллионов хэндлов не являются общедоступными. Система Handle поддерживается шестью глобальными корневыми сервисами, которые обрабатывают около 200 млн запросов в месяц, и более чем 2 тысячами локальных сервисов (LHS). К прокси-серверам CNRI происходит более 175 млн обращений в месяц. Сервисы DOA представлены в 75 странах мира (более тысячи сервисов).

Основные области применения архитектуры:

- Библиотеки и архивы;
- Интеллектуальная собственность;
- Дистанционное обучение и научные исследования;
- Интернет вещей;
- Индустрия развлечений;
- Контроль за цепочкой поставок и борьба с подделками.

Среди пользователей DOA такие известные национальные и международные организации как:

- Библиотека Конгресса США;
- Технический информационный центр Министерства обороны США;
- ОЭСР (OECD);
- Массачусетский технологический университет;
- Министерство сельского хозяйства США;
- CrossRef (ассоциация научных издателей, более 2 тысяч членов);
- Австралийское агентство по авторским правам;
- Nielsen BookData (ISBN);
- R.R. Bowker (ISBN);
- Национальная научно-техническая библиотека Германии;
- Национальная исследовательская лаборатория Лос Аламос;
- МСЭ (ООН) и многие другие.

Существует система <http://www.crossref.org> для регистрации научных и любых иных публицистических произведений, в которой зарегистрировано на сегодняшний день более 70 миллионов объектов. Использование идентификаторов DOI для присвоения публикациям стало де-факто стандартом в научном и академическом сообществе. В 2014 году Некоммерческое

партнерство «Национальный Электронно-Информационный Консорциум» (НЭИКОН) получило статус первого в России официального партнера и агента Международной ассоциации по связям издателей (Publishers International Linking Association, Inc. – PILA), корпорации, являющейся владельцем и оператором базы данных цифровых идентификаторов CrossRef. Таким образом, НЭИКОН оказывает содействие научным изданиям Российской Федерации в получении такого важного инструмента повышения цитируемости статей как Цифровой идентификатор объекта (DOI), от чего во многом зависит развитие российской науки. Некоторые русскоязычные журналы также начали публиковать DOI номера, например, журнал «Успехи физических наук» и некоторые другие. Свой уникальный DOI-номер получает как сам журнал, так и каждая статья в нем.

В 2015 году Фонд содействия развитию Интернета «Фонд поддержки интернет» стал членом организации International DOI Foundation (IDF) в статусе General Member (GM). С 2015 года основным направлением деятельности Фонда стало внедрение технологии Digital Object Identification (DOI) в России.

Deep Carbon Observatory (DCO)

Этот международный междисциплинарный проект по изучению углеродных циклов Земли является одной из первых реализаций реестра типов DOA. Научная сеть DCO состоит из более 400 организаций в 40 странах, в которых участвует более 1700 ученых. В рамках проекта создается концептуальная модель взаимодействия между данными, людьми, изданиями, инструментами, моделями, организациями, репозиториями и т.д. В процессе идентификации и аннотации всех ключевых сущностей, участников и типов активностей создается особый репозиторий данных и метаданных: инструменты для совместной работы ученых в области изучения Земли и космоса. Целью создания реестра DTR является идентификация различных структур представления научных данных (в форме реестра типов). Таким образом, в будущем можно будет установить нормы и правила курирования и дальнейшего использования данных в научной среде.

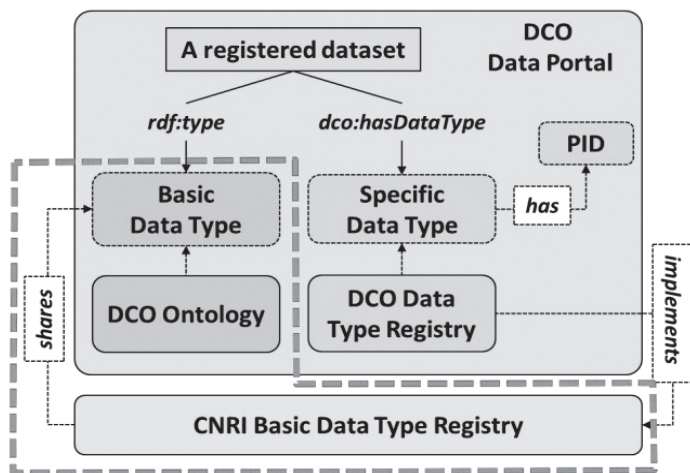


Рисунок 8. Прототип реестра типов DCO

В процессе аннотации набор данных сравнивается с базовыми наборами типов (изображение, аудио, видео, и тд.) в онтологии DCO, и ему присваивается один из этих базовых типов. Затем атрибуты метаданных сравниваются с реестром типов, зарегистрированных в DCO, и происходит связывание соответствующих атрибутов с идентификаторами в реестре.

Участие в подобных проектах российских исследователей помогло бы преодолеть языковой барьер для использования материалов научных работ, помогло бы развитию экспериментальной науки.

EIDR

Один из наиболее удачных примеров реализации реестра с использованием идентификации контента с помощью DOI демонстрирует EIDR (<http://eidr.org/>) – Entertainment Identifier Registry (Регистратура идентификаторов развлекательного контента), который смог воплотить на практике уникальную систему идентификации кинофильмов и телепередач. EIDR является глобальным реестром для уникальной идентификации для фильмов и телевизионного контента, нацеленным на автоматизированный обмен данными. Система EIDR была создана для того, чтобы сделать цифровую дистрибуцию

конкурентоспособной, повысить уровень взаимодействия и автоматизации между различными приложениями и платформами. Реестр позволяет учитывать всю широту иерархии кинопродукции, включая различные форматы и версии фильмов (включая физические носители), страны выхода и языки перевода, имеющиеся связи с другими сущностями в системе. EIDR реализован на основе DOI и включает публичные API и инструменты для разработки, доступные для членов организации. Членами EIDR являются около 100 крупнейших киностудий и производителей видеоконтента, в том числе 20th Century Fox, Walt Disney, Paramount Pictures, Warner Bros, HBO, MovieLabs, Netflix и др.

EIDR позволяет создавать несколько видов регистраций видеоконтента:

- Цикл, являющийся контейнером для сезонов и эпизодов;
- Сезон, являющийся контейнером для эпизодов;
- Эпизод, являющийся частью цикла и сезона.

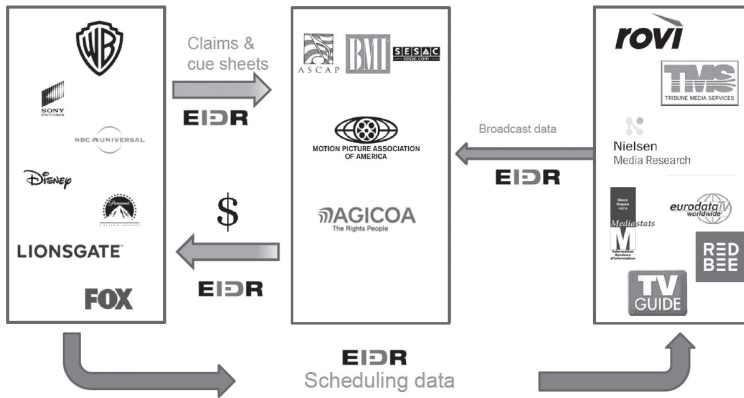


Рисунок 9. Участники системы EIDR

При этом система позволяет присваивать регистрируемым произведениям следующие метаданные: тип структуры, режим, референтный тип, наименование произведения, а также его альтернативное название, оригинальный язык, версия языка, продюсерская компания, дата релиза, страна производства, статус.

The screenshot shows the EIDR search interface with the following elements:

- Header:** EIDR Entertainment Identifier Registry logo, navigation tabs (Home, Search, Register), and a search bar (Content ID, Party ID) with a LOOKUP button.
- Navigation:** Content, Parties, Video Services tabs.
- Search Options:** Use ASCII search (checked), Results per page (25).
- Filters:**
 - Title, Language, Title Class, Referent Type, Registrant, Country of Origin.
 - Exact Title Match (unchecked), Include Alternate Titles (checked).
 - Associated Org Name, Associated Org ID, Role, Director, Actor, Actor #2.
 - Include Alternate Names (checked), Last Modified (From/To), Release Date (From/To).
 - Create Date (From/To), Alternate ID Type, Alternate ID, Alternate ID Relation.
 - Alternate ID Domain.
 - Inheritance Relationships, Root Objects Only (unchecked).
 - Lightweight Relationships.

Рисунок 10. Интерфейс EIDR

Таким образом, поиск необходимого произведения становится очень простым. Любой видеоконтент, зарегистрированный в системе, можно найти с помощью непосредственно DOI-идентификатора, названия произведения, наименования правообладателя и любых других критериев, необходимых для выборки самим пользователем.

Регистратура EIDR использует сложную систему дедупликации данных, чтобы убедиться в том, что объект, представленный для регистрации, ранее не регистрировался. Однако это не касается регистрации похожих или связанных с ним объектов.

Централизованная структура реестра гарантирует уникальность регистрируемых объектов. Если дубликата объекта в системе нет, регистратор в автоматическом режиме генерирует EIDR-код для объекта, сохраняет EIDR и соответствующие метаданные в реестре. При этом EIDR предоставляет набор API для разработчиков сторонних приложений и услуг, таких как интеграция производственных процессов, отчетность и связанность метаданных. Однако EIDR не является источником достоверных сведений о принадлежащих правах. Это задача не решалась в момент построения системы. EIDR представляет

собой уникальную систему для коммерческого взаимодействия между правообладателями контента и его пользователями в сегменте B2B.

Content ID	Title	Title Lang.	Ref. Type	Relationships	Original Lang.	Lang. Mode.	Structural Type	Release Date
10.5240/E1E8-4A1B-30CD-CE83-729C-H	Adventures Into Digital Comics	en	Movie		en	Audio	Abstraction	2006
10.5240/FB8B-DD15-4DAE-4574-5D90-7	Afro @Digital	en	Movie		en	Audio	Abstraction	2003
10.5240/BC3A-CA86-ACF2-B048-8802-D	Born Digital - Jugend 2.0 am Beispiel Prix Ars Electronica	de	TV		de	Audio	Abstraction	2011
10.5240/F1FB-8A30-5644-7706-067C-5	Cybersecurity: Defense in the Digital Age	en	TV	isEpisodeOf (10.5240/2948-2862-AA4F-D54A-2109-C)	en	Audio	Abstraction	2012
10.5240/E02C-7A82-3916-B08E-56B7-Q	Cyberstalker	en	Movie		en	Audio	Abstraction	1995
10.5240/7521-E7A1-8735-50AB-719E-L	Dealing with Skin Tones in Digital Photography	en	TV	isEpisodeOf (10.5240/1008-151B-E222-0C74-8AE0-U)	en	Audio	Abstraction	2006
10.5240/8E22-E46F-DABC-1B85-24F4-D	Digital Babylon	en	TV	isEpisodeOf (10.5240/8406-0830-88CE-CE83-DAF5-2)	en	Audio	Abstraction	2000

Рисунок 11. Интерфейс поиска в EIDR

С помощью EIDR возможна реализация таких сервисов, как автоматизированная доставка контента по требованию (VOD), динамическое размещение рекламы, кроссплатформенные изменения аудитории, управление релизными окнами, новые модели продажи контента. Не секрет, что правовые отношения в области продажи контента могут быть чрезвычайно запутанными. Точное описание метаданных объектов позволяет стандартизировать и автоматизировать процессы, связанные с электронной коммерцией вокруг объектов авторского права (продажа, лицензирование, и т.д.), поддерживая интероперабельность между операторами, продавцами, платформами и правообладателями в глобальном масштабе.

Пример описания модели данных EIDR⁵⁹

Для такого формата визуального контента, как сериалы, характерно разделение на сезоны, каждый из которых, в свою очередь, делится на эпизоды. Здесь приведено описание атрибутов

⁵⁹ Полное описание используемой модели данных занимает около 100 страниц и доступно по адресу: http://eidr.org/documents/EIDR_2.0_Data_Fields.pdf

модели данных для сериалов, а также специфических атрибутов для сезонов и эпизодов.

Общая модель данных для фильмов

В приведенной ниже таблице содержатся описание общих атрибутов для телевизионного контента:

Элемент	Описание
StructuralType	Обязательное поле. Требуемое значение этого поля – <i>Abstraction</i> .
Mode	Обязательное поле. Требуемое значение этого поля – <i>AudioVisual</i> .
ReferentType	Обязательное поле. Требуемое значение этого поля – <i>Series</i> .
ResourceName	Обязательное поле. Содержит оригинальное название сериала. Из соображений преемственности оригинальное название предпочтительно. EIDR. Для «экзотических» языков (в оригинальном тексте приводится японский в качестве примера, “Hagane no Renkinjutsushi” вместо 鋼の錬金術師) рекомендуется использования транслитерированного названия для удобства пользователей. Оригинальное название при этом может быть указано в качестве альтернативы. Обязательным атрибутом является язык (атрибут <i>lang</i>) со значениями в соответствии с ISO-кодами. Атрибут <i>titleClass</i> должен иметь значение <i>release</i> или <i>transliterated</i> . При регистрации сериала до того, как окончательно утверждено название, этот атрибут может иметь значения <i>working</i> или <i>internal</i> . В таком случае рекомендуется сразу, как только окончательное название становится известно, изменить название.
Alternate ResourceName	Необязательное поле. Включает в себя наименование на других языках и предназначено для облегчения идентификации. Также требует указания атрибута <i>lang</i> . Может содержать альтернативные имена с указанием специальных значений атрибута <i>titleClass</i> , таких как <i>AKA</i> , <i>acronym</i> или <i>other</i> .

OriginalLanguage	Обязательное поле. Язык, использованный актерами в процессе съемки, указанный в соответствии с IETF RFC5646. Может включать опциональный атрибут <i>mode</i> , который должен иметь значение <i>audio</i> для телевидения, и некоторые другие атрибуты.
AssociatedOrg	Обязательно, если не заполнено поле <i>Credits</i> . Должно содержать компанию-продюсера первоначальной трансляции. Должно быть указано минимум 1 значение, максимальное количество – 16. Может содержать ссылку на организацию, зарегистрированную в EIDR в качестве производителя контента. Обязательным является атрибут, указывающий на роль компании при создании сериала. Для компании-продюсера должно быть указано значение <i>producer</i> .
ReleaseDate	Обязательное поле. Год появления первого эпизода. Месяц и день могут быть добавлены, но не обязательны.
CountryOfOrigin	Обязательное поле. Страна, для которой изначально предназначался сериал.
Status	Обязательное поле. Допустимые значения – <i>valid</i> или <i>in development</i> .
ApproximateLength	Обязательное поле. Указывается примерная продолжительность эпизода.
AlternateID	Необязательное поле. Идентификатор в других системах, таких как IMDB.
Registrant	Ссылка на сторону, регистрирующую объект.
Credits	Поле является обязательным, если не указано поле <i>AssociatedOrg</i> . Может включать в себя до двух режиссеров и до 4 ведущих актеров. Для сериалов рекомендуется не указывать это значение, так как режиссеры и ведущие актеры могут меняться на протяжении съемок. Поле используется для проверки на уникальность регистрируемого объекта.
Description	Необязательное поле. Может содержать описание сериала.

Специфичные атрибуты сериалов

В приведенной ниже таблице приведены атрибуты, специфичные для сериалов как частного случая телевизионного контента:

Элемент	Описание
EndDate	Необязательное поле. Дата последнего эпизода в сериале. Должно отсутствовать в случае, если не известна дата окончания. Рекомендуется для валидации дат, указанных для отдельных эпизодов.
SeriesClass	Необязательное поле. Обычно имеет значение <i>Episodic</i> , другие возможные значения – <i>Anthology</i> или <i>Mini-series</i> .
NumberRequired	Необязательное поле.
DateRequired	Необязательное поле. Обычно имеет значение <i>false</i> , если сериал включает в себя сезоны.
OriginalTitle-Required	Необязательное поле. Обычно имеет значение <i>false</i> , если сериал включает в себя сезоны, так как у сезонов нет собственных заголовков.

Специфичные атрибуты сезонов

В приведенной ниже таблице приведены атрибуты, специфичные для отдельных сезонов:

Элемент	Описание
Parent	Обязательное поле. Идентификатор сериала.
EndDate	Необязательное поле. Дата последнего эпизода сезона. Не должна быть указана для текущего сезона или в случае, когда дата окончания неизвестна. Рекомендуется указывать для валидации дат отдельных эпизодов.
SeasonClass	Необязательное поле. Обычно имеет значение <i>Main</i> , может также иметь значение <i>Mini-series</i> .
Number-Required	Необязательное поле, рекомендованное к использованию. Если эпизоды сериала упорядочены, должно иметь значение <i>true</i> .
DateRequired	Необязательное поле. Обычно имеет значение <i>true</i> для ежедневных программ или в случае, когда отсутствуют другие идентификации, такие как дата и название серии.
OriginalTitle-Required	Необязательное поле. Рекомендованное значение для современных художественных фильмов – <i>true</i> .
Sequence-Number	Необязательное поле. Целочисленное значение, идентифицирующее номер сезона. Первый сезон должен иметь номер 1.

Специфические атрибуты отдельных эпизодов

В приведенной ниже таблице приведены атрибуты, специфичные для отдельных эпизодов:

Элемент	Описание
Parent	Обязательное поле. Должно содержать идентификатор сериала или сезона сериала.
EpisodeClass	Необязательное поле. Обычное значение <i>Main</i> , другие значения – <i>Pilot</i> , <i>Standalone</i> , <i>Special</i> , <i>Omnibus</i> , <i>Recut</i> , <i>Segment</i> .
Distribution-Number	Если у сериала или сезона, указанного в качестве родительского объекта, установлено в <i>true</i> свойство <i>NumberRequired</i> , то поле является обязательным. Рекомендуется указывать поле в случаях, когда оно известно. Поле обозначает номер эпизода в рамках родительского объекта, основываясь на дате показа. Нумерация начинается с 1 для каждого сезона. Требуется, чтобы номер каждого эпизода в рамках родительского объекта был уникальным. Возможна более сложная логика нумерации. Может содержать опциональный атрибут <i>domain</i> .
HouseSequence	Необязательное поле. Предназначено для указания номера эпизода до того, как окончательная идентификация стала известна. Может содержать опциональный атрибут <i>domain</i> .
AlternateNumber	Необязательное поле. Может содержать до 32 альтернативных номеров эпизодов. Должна содержать атрибут <i>domain</i> .

Реализация DOA в Китае

Одни из наиболее ярких примеров DOA реализуются в Китае. Так, с 2006 года при поддержке компании HP и Министерства образования Китая реализуется проект China Digital Museum Project, нацеленный на создание цифрового музея, объединяющего коллекции более ста университетов по всей стране.

С 2007 по 2010 год Министерство науки и технологий финансировало международный исследовательский проект по разработке и применению решений на основе системы уникальных идентификаторов цифровых объектов совместно с CNRI. На его основе, в том числе, была создана цифровая система управления

авторскими правами, использующая Handle и DOI для регистрации и проверки прав на контент.

В последние годы ETIRI работает над двумя национальными проектами, где DOA и Handle выбраны в качестве технологической основы – это Национальная платформа по отслеживанию качества пищевой продукции (также используется и для лекарственных препаратов) и Национальная общественная платформа для управления идентификацией в интернете вещей.

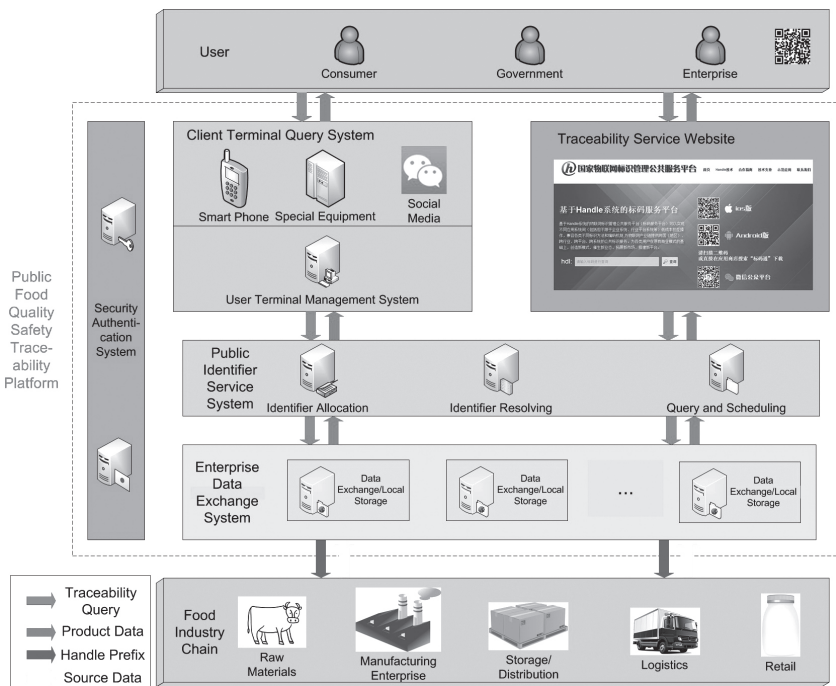


Рисунок 12. Китайская система отслеживания качества продукции на основе DOA

Качество продукции отслеживается на каждом этапе индустриальной и логистической цепочки, при этом данные из каждого источника сохраняются в системе для каждой произведенной единицы продукции. Соответствие информации на любом этапе можно проверить с использованием встроенных QR-кодов или RFID-чипов. На начало 2015 года с помощью системы было

отслежено около 100 млн единиц продукции от шести крупнейших производителей.

В 2014 году на конференции МСЭ в Бусане была принята резолюция по борьбе с контрафактными технологическими устройствами с использованием рекомендаций ITU-T X.1255, которая основывается на архитектуре DOA. Общая схема работы такова:

- Производитель регистрирует хэндл для каждого производимого устройства;
- Для каждого устройства генерируется и сохраняется в метаданных цифровая подпись;
- Информация о поставках добавляется к объекту по мере прохождения логистической цепочки от производства к складированию, от дистрибуторов к ритейлерам;
- Во время покупки потребитель получает необходимую информацию об устройстве: цифровую подпись, уникальные идентификаторы устройства (IMEI, MAC, серийный номер и т.д.), информацию о предполагаемом месте продажи товара;
- Сравнивая информацию с содержащейся в системе (используя QR-код или RFID для считывания хэндла), покупатель может убедиться в подлинности устройства.

Возможности использования реестра цифровых объектов в России

В России неоднократно звучали предложения по созданию реестра объектов авторского права, как в рамках работы Министерства связи и массовых коммуникаций, так и со стороны правообладателей и обществ по управлению коллективными правами. Создание такого реестра с использованием технологии DOA/DOI для регистрации произведений, компиляций, наборов данных и других объектов авторских прав обеспечило бы решение ряда важных на сегодня задач:

- Однозначную идентификацию контента для широкого круга пользователей, правоприменительных органов и судов;
- Идентификацию авторов и правообладателей контента;
- Уведомление широкого круга пользователей о выбранном объеме защищаемых прав на контент;
- Снижение рисков нарушения прав на контент;

- Повышение совместимости активов и метаданных из различных баз данных, сервис-провайдеров, поставщиков метаданных;
- Более полное представление контента и связанных с ним объектов в интернете.

Использование систем идентификации товаров для проверки их подлинности и последующей продажи как в онлайн-ритейле, так и в офлайн-магазинах представляется одной из наиболее перспективных областей внедрения решений на основе DOA в настоящий момент. В России существуют такие обязательные для участников рынка системы как ЕГАИС; существуют предложения по созданию аналогичных систем учета продукции в области фармацевтических препаратов, пищевой промышленности, в системе государственного кадастрового учета и другие. При этом каждый такой проект сопровождается многочисленной критикой работы внедряемых аппаратно-программных решений, дороговизна разработки, ненадежность и неэффективность систем. Использование решений на основе архитектуры DOA, по аналогии с китайским опытом, могло бы сократить расходы государства, дать стимул реальному сектору экономики и повысить проникновение ИКТ в тех областях, в которых будут внедряться подобные решения.

Создание единого и удобного средства обеспечения гарантированной идентификации и аутентификации для физических и юридических лиц в России для получения государственных, муниципальных услуг и информации, развития юридически значимой электронной коммерции, в том числе финансовых операций, а также повышения степени ответственности участников информационного обмена входит в дорожные карты Института Развития Интернета, создаваемые на основе поручения Президента РФ. Модернизация работы судебной системы путем внедрения электронного нотариата и судопроизводства также является частью Программы развития интернета в России. Хорошим шагом на пути реализации подобных стратегических документов было бы обязательное требование к изучению возможности применения существующих стандартов идентификации цифровых объектов и системы DOI/DOA в частности.

Система CrossRef (<http://www.crossref.org>) для регистрации научных и любых иных публицистических произведений, в которой зарегистрировано на сегодняшний день более 70 млн. объектов, а также система EIDR для регистрации видеоконтента, в которой зарегистрировано более 700 тыс. объектов, могут использоваться в качестве модельных примеров при построении национального реестра произведений.

Различные виды контента могут регистрироваться разными регистраторами в зависимости от типа самого контента. Каждый регистратор может проводить свою политику, не противоречащую принципам технологии DOA/DOI, осуществляющей управление системой и отвечающей за стандартизацию системы. Это обеспечивает регистрацию и идентификацию контента с использованием единообразного алгоритма “Handle system”, являющегося сердцем системы. Кроме того, важнейшей задачей, стоящей перед регистратором, является обеспечение уникальности регистрации объекта, т.е. недопущения как повторения идентификаторов, так и повторной регистрации объекта в системе с другим, пусть и уникальным, идентификатором.

Реестр должен состоять из записей, вносимых регистратором, который получает и обрабатывает запросы на регистрацию. Сами запросы направляются регистрантами (администраторами), а пользователи и приложения могут просматривать и искать информацию в реестре. При этом регистранты и пользователи могут получить доступ к системе либо с помощью веб-интерфейса самой системы, либо с использованием API для технологического взаимодействия с реестром.

Регистрантом контента может выступать как сам владелец контента, так и агрегатор, компания пост-продакшн либо иное лицо, авторизованное на регистрацию объектов.

В отношении музыкальных произведений в реестре могут быть зарегистрированы как сами объекты, так и производные объекты: компиляции, дискографии, альбомы, мини-альбомы (EP) и отдельные треки. При этом в реестр могут быть внесены следующие метаданные о произведении: наименование, жанр, дата звукозаписи, дата релиза, язык, автор музыки и текста,

студия звукозаписи, наименование артиста, объем авторских прав на произведение, цифровой отпечаток и др.

Для книг это может быть название книги, жанр, автор(ы), год выпуска в свет, издательство, ISBN, УДК, ББК и другие используемые идентификаторы.

Для видеоконтента за основу может быть принята уже реализованная система EIDR, речь о которой шла выше.

Создание экосистемы, основанной на DOA/DOI, позволит в дальнейшем использовать публичный реестр произведений и зарегистрированные имена DOI для любого вида управления контентом – как коммерческого, так и некоммерческого. Реестр, функционирующей на архитектуре DOA/DOI – это мультизадачное решение по множеству актуальных вопросов для большого количества участников рынка оборота цифрового контента, так как набор метаданных может включать не только характеристики произведений и информацию о правообладателе, но также и содержать сведения об объеме исключительных прав на произведения. Это в свою очередь позволит идентифицировать не только произведения, защищенные авторским правом в полном объеме, но и произведения, распространяемые на условиях свободных (открытых) лицензий, а также те произведения, которые уже перешли в режим общественного достояния, однако по причине отсутствия технического решения не были надлежащим образом учтены и каталогизированы.

Стоит также отметить, что в рамках системы DOA/DOI могут быть идентифицированы не только произведения, изначально созданные в цифровой форме, но также и любые иные произведения, которые уже существовали в аналоговой форме и позже были оцифрованы.

В последнее время большое внимание уделяется вопросу создания «электронного государства». Под данным понятием, как правило, определяется особый комплекс организационно-технических средств взаимодействия государства с гражданами, бизнесом, чиновниками и отдельными органами власти. «Электронное государство» – это политико-правовая организация общества, представляющая собой способ взаимодействия

органов государственной и муниципальной власти и граждан в целях предоставления государственных услуг и обеспечения возможности участия в осуществлении власти населения с использованием информационно-телекоммуникационных технологий, обеспечивающая его единство и территориальную целостность, в том числе в электронном пространстве, осуществляющая посредством государственного электронного механизма управление делами общества, суверенную публичную власть, придающую праву общеобязательное значение, гарантирующую права, свободы граждан, законность и правопорядок.

В рамках «электронного государства» традиционное бумажное делопроизводство и взаимодействие начинает заменяться широким и эффективным использованием сетевых компьютерных технологий и сопутствующих программных средств.

Основным двигателем взаимодействия между государством и обществом являются развивающиеся ИТ-технологии. В рамках указанной концепции первостепенное значение играет не столько нормативно-правовая база, которая, безусловно, должна соответствовать складывающимся в этой сфере отношениям и не препятствовать их развитию, сколько реальные действия исполнительной власти по внедрению высоких технологий.

На сегодняшний день разработана и функционирует электронная система разрешения споров, которая уже успешно применяется в практиках некоторых западных стран. Стокгольмская программа ЕС (2010/С 115/01) от 4 мая 2010 года⁶⁰ настоятельно рекомендует использовать «электронное правосудие».

Интернет-судопроизводство основано на широком использовании электронных документов. Вместе с тем главной проблемой электронных документов является проблема юридической силы подобных документов.

Представляется, что система DOI могла бы стать эффективным решением в рамках электронного нотариата и электронного судопроизводства, обеспечивая при этом достоверность

60 The Stockholm Programme – An open and secure Europe serving and protecting citizens. 2010 / С 115/01. Опубликовано в официальном журнале NC 115, 4.5.2010.

информации, указание на местонахождение информации, а также однозначную идентификацию субъекта, являющегося источником информации. С помощью архитектуры “цифровых идентификаторов” субъекты гражданско-правовых отношений имели бы возможность точно идентифицировать контент в целях его дальнейшего оборота.

Формирование “электронного государства” в России также является необходимым шагом для совершенствования взаимодействия между обществом и властью, эффективного управления, искоренения бумажной волокиты во всех государственных структурах, в том числе и судах. Построение инфраструктуры для создания механизмов “электронного государства” должно стать приоритетной задачей на федеральном уровне. Без ее реализации укрепление позиций России на международной арене в сфере информационно-коммуникативных технологий и электронного государственного управления невозможно.

Использование цифровых доказательств является достаточно актуальной проблемой для судопроизводства. Даже несмотря на то, что ст. 58 ГПК содержит такую процедуру представления доказательств как осмотр и исследование письменных или вещественных доказательств по месту их хранения или месту их нахождения, суды неохотно применяют эту возможность на деле, когда речь идет о спорах, связанных с размещением информации в сети Интернет. Вместо непосредственного изучения страниц сайтов в сети Интернет, суды требуют от участников процесса заверения доказательств у нотариуса и представления скриншотов в распечатанном и заверенном виде, либо привлечения экспертов, которые смогут подтвердить или опровергнуть наличие или отсутствие информации в сети Интернет.

Очевидно, что в эпоху развитых цифровых технологий все это вызывает большое число сложностей у участников процесса ввиду дороговизны процедуры заверения цифровых доказательств и больших временных затрат, необходимых для нотариального заверения страниц сайтов в целях обеспечения достоверности представляемых суду доказательств. Кроме того, это непременно ведет к судебной волоките.

В ходе совершения нотариального действия по заверению страниц сайтов как доказательств по гражданским делам нотариус осматривает соответствующий информационный ресурс начиная с главной (стартовой) страницы веб-сайта, распечатывает его содержимое на бумажный носитель и составляет протокол осмотра доказательств. Для предотвращения возможных фальсификаций на практике в процессе осмотра доказательств в интернете нотариусы зачастую применяют специальные программы-утилиты для повышения достоверности полученных данных, в т.ч. сервис WhoIS, позволяющий определить администратора доменного имени, под которым размещен осматриваемый веб-сайт; программу nlookup, позволяющую определить IP-адрес веб-сайта с использованием данных, полученных с помощью сервиса Whois; программу ping, позволяющую проверить, совпадает ли определенный посредством программы nlookup IP-адрес с тем, к которому обращается браузер; а также программу tracert, с помощью которой отслеживается маршрут доступа к целевому веб-сайту

Однако следует отметить, что нотариус не всегда имеет право заверять доказательства. Согласно ч. 2 ст. 102 Основ законодательства о нотариате, не допускается возможность обеспечения нотариусом доказательств по делам, уже находящимся в производстве суда. Данная норма является следствием соблюдения принципа непосредственности судебного разбирательства, закрепленного в ч. 1 ст. 10 АПК РФ, в соответствии с которой арбитражный суд при разбирательстве дела обязан непосредственно исследовать все доказательства по делу, а также следствием наличия процедуры обеспечения доказательств непосредственно в самом ГПК РФ.

Система DOA, которая однозначно определяет сам контент, путь к нему или иному цифровому объекту в общем информационно-виртуальном пространстве, а также правообладателя контента, могла бы стать отличным решением проблемы идентификации, изучения и оценки цифровых доказательств.

DOA способна решить также и важную задачу для онлайн сервисов, которые, используя систему цифровых идентификаторов,

получили бы возможность однозначно идентифицировать все свои правообразующие документы (в том числе учредительные документы, правила пользования, общие условия предоставления услуг/выполнения работ, положения о конфиденциальности и об ограничении ответственности, локальные политики сервисов по обработке персональных данных и др.) и юридически значимые сообщения (заявления, уведомления, извещения, требования) для дальнейшего доказывания своей правовой позиции в рамках административных и судебных споров.

Управление и супервайзинг соблюдения стандарта авторитетной, независимой от сторон спора некоммерческой организацией, а также автоматизированный процесс присвоения уникальных цифровых идентификаторов, исключающий возможность манипуляций с датой и субъектом регистрации, способны гарантировать высокий уровень доверия со стороны судов и правоприменительных органов к такого рода цифровым доказательствам.

Увеличение «электронной составляющей» гражданского оборота и применение электронных технологий в международном нотариальном праве подчеркивается также и Международным союзом нотариата. Согласно уставу Международного союза нотариата использование современных технологий способно сократить сроки исполнения административных формальностей.

Электронная нотариализация позволяет обеспечить юридическую силу и достоверность документов с помощью проставления на электронном документе «метки времени», выполнения нотариального действия в электронном виде, а также электронной подписи правообладателя информации. Метка времени является доказательством того, что документ существовал в определенный момент, и позволяет проследить вносимые в документ изменения. Проставление «метки времени» необходимо, чтобы впоследствии невозможно было бесследно изменить содержание электронного нотариального документа.

Очевидно, не все юридические документы могут быть представлены в публичном доступе для широкого круга лиц. Документы, составляющие банковскую, налоговую, коммерческую тайну, а

также иные документы ограниченного распространения могут быть расположены в закрытых репозиториях, ведение которых может быть возложено на уполномоченные органы и нотариусов. Так, например, согласно ст.1123 ГК РФ, предусматривающей соблюдение тайны завещания, лица, имеющие доступ к сведениям, содержащимся в единой информационной системе нотариата, и лица, осуществляющие обработку данных единой информационной системы нотариата, а также гражданин, подписывающий завещание вместо завещателя, не вправе до открытия наследства разглашать сведения, касающиеся содержания завещания, его совершения, изменения или отмены.

Не является разглашением тайны завещания представление нотариусом, другим удостоверяющим завещание лицом сведений об удостоверении завещания, отмене завещания в единую информационную систему нотариата.

В случае внедрения системы DOA в единую информационную систему нотариата, нотариус, заверяющий завещание, имел бы возможность инициировать присвоение такому завещанию уникального идентификатора с указанием в метаданных места расположения самого документа в виде цифрового объекта в репозитории. Кроме того, в метаданных, соответствующих идентифицированному объекту, будут находиться сведения о времени, месте, участниках создания и хранения, ответственных лицах и прочие сведения. Технология DOA позволяет регламентировать доступ к конкретным метаданным, в том числе с помощью системы разграничения доступа (системы управления полномочиями), а также шифрования контента. В дальнейшем все заинтересованные лица (суд, Росреестр, банки, ГИБДД) по DOI-идентификатору могли бы получать сведения о факте наличия завещания в информационной системе нотариата с учетом своих полномочий при помощи технологии аутентификации и с использованием публичных ключей. В определенных случаях ограниченному кругу авторизованных пользователей может быть даже предоставлен доступ к тексту завещания для совершения юридически значимых действий.

Раздел 3.

Современные инструменты регистрации и идентификации в сети Интернет произведений в сфере культуры, науки и образования

Установление основных категорий произведений культуры, науки и образования (аудио, видео, текст, изображение и др.) по форматам файлов и целям использования

Материальные объекты и их свойства

Понятие материального объекта является в высшей степени абстрактным философским понятием, т.н. философской категорией. Спецификой философских категорий является их низкая содержательность – количество атрибутов, т.е. существенных признаков, входящих в определение категории, предельно невелико. Следует иметь в виду, что в академической среде не существует согласия относительно того, какие определения являются корректными или приоритетными, т.к. это зависит от концептуальных установок тех или иных теоретических доктрин. Применительно к задаче настоящего исследования это означает, что нет большого практического смысла формулировать определение материального объекта. Гораздо более осмысленным является подход, согласно которому дается некое описание признаков (существенных и несущественных) того, что понимается под таковыми, и классификация материальных объектов.

Признаки материальных объектов можно разделить на следующие классы по предметному критерию:

- 1) пространственные характеристики;
- 2) темпоральные свойства, т.е. время существование объекта;
- 3) физические свойства, (применительно к материальному предмету – свойства твердых тел);
- 4) химические свойства (химический состав);
- 5) физико-химические свойства (электро-химические и радиологические характеристики).

С точки зрения критерия научных инструментариев, применяемых для выявления перечисленных классов свойств, все материальные свойства объектов устанавливаются исключительно с помощью *физических* и *химических* методов. *Это означает, что никаких иных объективных методов, кроме методов физики и химии для идентификации материальных объектов не существует.*

1) Пространственная протяженность. Это означает, что материальным объектам можно приписать различные пространственные метрические характеристики на основании измерений. Такие характеристики делятся на абсолютные и относительные:

А) *Абсолютными метрическими* характеристиками являются геометрические характеристики, позволяющие идентифицировать абсолютные размеры материального объекта согласно принятыми системами мер. Они выражаются в таких классических параметрах евклидовой геометрии, как длина граней, диагональная длина, ширина, высота, глубина (полостей), диаметр (внешний и внутренний), радиус. В инженерных дисциплинах для удобства евклидовые параметры могут быть конкретизированы в виде таких технических параметров как калибр (внутреннего и внешнего) шага винта, количества и шага нарезки и т.п.

Б) *Относительные метрические* характеристики используются для двух целей: 1) определения формы объекта, безотносительно к его абсолютным размерам (позволяют установить наличие/отсутствие отношения подобия между сходными объектами); 2) определения местоположения стационарного объекта (например, объекта недвижимости) относительно геодезических систем координат и относительно других объектов. Такие

характеристики обычно выражаются в параметрах сферической или афинной геометрии. К таким параметрам относят угловые параметры, параболические параметры, координаты объекта в топографической и сферической системе координат.

2) Темпоральные характеристики. Это время существования материального объекта и точки его модификации во времени (переделки, достройки, реставрации, порчи, изготовление палимпсеста и т.п.). Следует иметь в виду, что точно установить темпоральные характеристики обычно исключительно сложно в отсутствии соответствующего документирования судьбы материального объекта.

2) Физические свойства твердых тел. Они могут быть абсолютными и относительными по своей математической форме, однако и те и другие обычно могут быть использованы для идентификации физического объекта с равным успехом. Эти характеристики обычно выступают в виде следующих параметров: масса, плотность, твердость, отражающая способность (альбедо), спектральные характеристики отраженного от объекта света в стандартных условиях, электропроводимость, магнитная ориентация (если наличествует), напряженность конструктивных элементов и т.п.

3) Химический состав вещества. Химические характеристики устанавливаются методами химического анализа. К ним обычно относят относительный химический состав, выражающийся в процентном отношении элементов периодической системы и более сложных химических веществ, входящих в состав материи объекта,

4) Физико-химические свойства. К данному классу свойств, применительно к предмету настоящей работы, в первую очередь следует отнести радиохимические свойства, используемые при радиоизотопной датировке объектов (радиоуглеродной, радиофосфорной и т.п.) и, возможно, различные электрохимические характеристики вещества объекта.

Следует иметь в виду, что ни одно из перечисленных свойств не является достаточным для идентификации материального объекта. Известно, что материальные объекты порой весьма успешно

копируются и подделываются. Однако, комбинация методов позволяет установить подлинность того или иного объекта в случае известности критериев, т.е. свойств подлинного объекта.

Материальные объекты обладают характерным теоретико-познавательным свойством. Знание о них *всегда неполно* – по той простой причине, что наше знание о физической реальности не является полным. Количество атрибутов и свойств материальных объектов потенциально бесконечно, поскольку обусловлено сложнейшим комплексом естественных причин, в том числе случайных. Появление новых физических и химических методов анализа позволяет выявлять новые свойства. Отличительным свойством материальных объектов по сравнению с цифровыми является то, что они обладают абсолютными физическими и химическими свойствами – их невозможно изменить, не «испортив» объекта, т.е. их изменение влечет за собой потерю объектом своего тождества – объект перестает быть самим собой. Материальный объект – это всегда конкретная вещь. Материальный объект обладает уникальным неконцептуальным содержанием, т.е., несмотря на то, что его можно определить с помощью общих понятий, подобное определение нельзя будет использовать для отличия подлинника объекта, например, объекта культуры, от его копий.

Цифровые объекты и их свойства

Содержание понятия *цифрового объекта* является не менее полемичным, чем содержание понятия объекта материального. Свойства цифровых объектов являются относительными по своей природе, в том смысле, что в качестве свойств объектов этого типа выступают логические отношения. Это означает, что свойства цифровых объектов описываются двухместными предикатами, в отличие от материальных объектов, свойства которых описываются одноместными предикатами. И в этом смысле свойства цифровых объектов являются реляционными по своей природе.

Цифровые объекты не обладают неконцептуальным содержанием. Цифровой объект не является конкретным объектом вообще, поскольку это класс логических конструкций.

Сравнительный анализ материальных и цифровых объектов.

Данный анализ представлен в таблице свойств материальных и цифровых объектов.

Материальные объекты	Цифровые объекты
Конкретное бытие, не сводимое к принадлежности к классу.	Абстрактное существование во внутренней форме упорядоченного множества и внешней – множества копий.
Наличие неконцептуального содержания. Например, для знания, что материальный объект S обладает свойством P, не обязательно владеть понятиями S и P.	Отсутствие неконцептуального содержания. Для полноты знания, чем является цифровой объект, достаточно знать имя и атрибуты объекта.
Потенциально неполное знание – знание о материальном объекте может быть достаточным с прагматической точки зрения в конкретных обстоятельствах, но всегда потенциально неполно.	Потенциально полное знание – знание о цифровом объекте полностью исчерпывается свойствами, заложенными в него при его создании. Оно может быть актуально бесконечным, т.к. цифровые объекты могут быть исключительно сложными для человеческого понимания, но потенциально оно конечно.
Принципиально неполноценная воспроизводимость («невозможность» создания принципиально недетектируемой копии уникального объекта)	Принципиально полноценная воспроизводимость (возможность создания полноценной копии цифрового объекта)
Локальность во времени и пространстве – всякий материальный объект в определенное время t пребывает в некотором месте, определяемом с той или иной точностью.	Нелокальность во времени пространстве – локальными во времени и пространстве могут быть только репрезентации цифровых объектов (копий), т.к. обычно можно сказать, где хранится копия цифрового объекта. Хранение данных в «облаках» размывает локальность данных даже конкретных репрезентаций цифровых объектов.
Уникальность (уникальный набор свойств)	Полная типичность, набор свойств полностью определяется логическими средствами.

Прибавочная стоимость является функцией от количества производимых объектов	Прибавочная стоимость почти не зависит от количества производимых объектов (ничтожно мала).
Личное отношение людей, анонимизация и сакрализация объектов.	Нет личного отношения к репрезентациям цифрового объекта. Копии не различимы и не формируют личного отношения.
Связанность объектов (система вещей) носит логический и субъективный ассоциативный характер. Действует в пределах человеческого разума или сети разумов (в общественном сознании). Степень их связности ограничена выразительными средствами подписей, этикеток, табличек, инвентарных журналов, бумажных реестров и человеческой памятью.	Связанность носит исключительно объективный логический характер как система гиперссылок одних объектов на другие. Любой объект при необходимости может содержать ссылку на любой другой объект.

Анализ положительных и отрицательных сторон цифровизации

Цифровизация или переход (поворот) внимания общества от материальных носителей культурного содержания в пользу их цифровых аналогов и репрезентаций является прогрессирующей тенденцией современной культуры. Эта тенденция наглядно проявляется в таких феноменах, как:

1) Рост количества оцифрованных произведений, в т.ч. знаний и культурных ценностей, включая постоянное увеличение количества цифровых объектов, в т.ч. в открытом доступе.

2) Рост популярности электронных «читалок» и, соответственно, книг в различных текстовых форматах для разных типов устройств.

3) Появление виртуальных музеев – интерактивных путеводителей по реальным достопримечательностям.

4) Рост количества вебкамер, транслирующих происходящее из различных мест: зоопарков, архитектурных комплексов и т.п.

5) Усиление тенденции гипертекстуальности – рост числа перекрестных ссылок с одних цифровых носителей на другие:

электронные энциклопедии, справочники, цифровые фотографии, аудио- и видеозаписи.

6) Рост доли цифровых технологий в создании нового культурного содержания: компьютерная анимация, электронная музыка, развитие наукоемкого искусства, использующего информационные технологии неожиданным образом.

7) Исследователи отмечают также такие сопутствующие цифровизации явления как переход от традиционных индустриальных моделей информации к новым коммуникативным практикам, предполагающим значительное повышение объема и качества использования цифровых объектов, правомерно размещенных в открытом доступе, т.к. они используются не только как источники информации, но и в коммуникации в т.ч. как символы и знаки. Этот эффект для объектов с неопределенным правовым статусом или неправомерных, к сожалению, уравнивается рисками для ресурсов и пользователей, а также связанными с ними ограничениями на свободу их использования.

Положительные стороны цифровизации:

Оцифровке подлежит большинство объектов материального мира, начиная от аудио и видеозаписей, заканчивая 3-D моделями материальных объектов. Существенное различие между аудио и видео объектами, представленными в сети, и оцифрованными объектами (текст, изображения) заключается в том, что при воспроизведении аудио/видео файлов пользователь может полноценно осуществить акт восприятия таких объектов. Оцифрованные изображения позволяют увидеть копию объекта материального мира и максимально возможно предположить о том, как он выглядит в реальном мире. Например, оцифрованная картина из галереи. Созерцая оцифрованное изображение художественного шедевра, мы можем лишь только представить, как он выглядит в реальном мире. За счет статуса копии, у оцифрованного изображения галерейной картины намного меньше материальная ценность, чем у оригинала, который был оцифрован. При демонстрации товаров в интернет-каталогах мы не имеем возможности представить пробный образец товара в сети таким образом, чтобы потребитель имел возможность

тактильно с ним взаимодействовать, в данном случае возможна только визуализация образа исходного объекта материального мира. Основываясь на представленном изображении товара, потребитель решает, имеет ли для него смысл отправляться в место его непосредственного хранения или же заказывать товар через интернет.

В то же время нельзя отрицать, что в ряде случаев цифровые объекты никак не уступают реальным объектам: например, качественная копия картины или цифровой текст могут вызывать те же эмоции, идеи и переживания, что и реальные объекты, превосходя их в потенциале как объекты коммуникации с точки зрения возможности такого копирования.

Таким образом, процесс цифровизации можно конкретизировать через следующие последствия:

1. Облегчение доступа к культурному содержанию вследствие многопорядкового удешевления издержек на копирования текстов, репродукций, фильмов и т.п.

2. Облегчение доступа к культурному содержанию жителей удаленных от культурных центров районов.

3. Рост связности культурного содержания как следствие гипертекстуальности – пользователь способен мгновенно получить огромный объем сведений о том или ином культурном объекте: авторе, культурно-исторических условиях создания, обусловленности другими объектами культуры, аналогичными объектами, роли объекта в истории, науке, политике и экономике.

4. Возможность обрабатывать статистику больших данных, собираемых с ПК и мобильных устройств для эффективного управления сферой культуры, науки и образования, коррекции культурной и образовательной политики.

5. Цифровизация является относительно дешевым и надежным способом сохранения культурного содержания для будущих поколений в случае утраты или гибели материального оригинала.

6. Независимо от того, насколько полноценно объект может быть представлен пользователю сети, оцифровывание объектов представляется необходимым этапом развития содержательного

компонента сети. Упрощение потребительского поиска, когда появляется возможность принять решение о необходимости приобретения товара при помощи первичной демонстрации изображаемого объекта.

Отрицательные стороны цифровизации:

1. Рост количества «нарушений прав на интеллектуальную собственность». За счет широкого доступа к информационной сети и постоянному пополнению сети разнообразным контентом может осуществляться незаконное копирование аудио/видео/текстовых файлов. Это связано с тем, что принятые в доцифровую эпоху схемы управления авторскими правами не очень хорошо работают в цифровую эпоху. Это не отрицательная сторона цифровизации, а несоответствие устаревших правил новой технологии.

2. Смена способа работы со знаковыми системами в массовом масштабе может иметь в последствии изменения когнитивных структур мозга, что необходимо учитывать при разработке стратегий образовательной политики. Эти изменения описываются в форме таких феноменов как «клиповое мышление», трудность восприятия больших текстов, неспособность эффективно пользоваться справочной литературой на бумажных носителях, недоразвитость моторных навыков вследствие депривации движения и минимизации необходимости писать от руки и т.п.

3. Снижение мотивации к заучиванию и запоминанию содержания вследствие возможности получить доступ к необходимой информации со своего мобильного устройства практически в любой момент времени.

4. Издательская деятельность является достаточно дорогостоящим видом деятельности, который могут себе позволить профессиональные институты – издательства, научные, образовательные учреждения и т.п. Это является определенного рода фильтром против информации низкого качества и ложных сведений. Радикальное удешевление генерации и компиляции культурного содержания, доступное для любого пользователя, независимо от его уровня компетенции, засоряет интернет недостоверной информацией, чем, к сожалению, активно пользуются

студенты, еще не обладающими достаточными навыками критического мышления.

Следует отметить, что следствия цифровизации не носят критического характера, не являются основанием для вывода о вредности стратегии цифровизации. Это неизбежные эпифеномены смены способов работы человечества со знаковыми системами в глобальном масштабе – «детские болезни» и неустраняемая плата за прогресс. Эти тенденции следует иметь в виду и вовремя демпфировать соответствующими мерами по:

1) переводу знаний и культурных ценностей, насколько это возможно, в режим открытого доступа или свободного обращения, если невозможно или противопоказано – в режим защиты информации от несанкционированного доступа;

2) сохранению и стимулированию традиционных форм туризма, подразумевающих посещение музеев, достопримечательных мест и природных ландшафтов;

3) соблюдению баланса между использованием компьютерной техники и иных способов доступа к культурному и научному содержанию в целях сохранения здоровья и моторных навыков;

4) стимулированию психофизических исследований, направленных на выявление изменений в головном мозге вследствие цифровизации образования с целью корректировки образовательной политики.

Отношение материальных объектов к цифровым. Проблема референции

Проблема референции является одной из фундаментальных философских проблем. Суть ее составляет спор о том, что связывает, например, высказывание «Это яблоко красное» с конкретным яблоком красного цвета. Существует немалое количество различных концепций референции. Применительно к нашей задаче нет смысла анализировать их подробно, однако, следует иметь в виду, что не существует однозначного необходимого и достаточного критерия, определяющего референциальное отношение репрезентации (высказывания, изображения) к материальному объекту как формы репрезентации последнего. Тем не менее, существуют устоявшиеся подходы в научной

и юридической практике, позволяющие осуществить подобную привязку. В практике делового оборота обычно это делается с помощью документирования репрезентации в форме инвентарной описи, накладной, сопроводительных надписей к фотографиям и кинозаписям, табличек к предметам и т.п. В случае отсутствия таковых, вывод о том, что то или иное изображение действительно является изображением конкретного объекта, делается экспертами на основании анализа большой совокупности эмпирических данных и теоретических знаний. Пример такой деятельности мы можем наблюдать в различных видах экспертиз, имеющих юридическое значение: искусствоведческой, графологической, автороведческой и т.п.

Цифровые объекты могут быть идентифицированы как конкретные репрезентации материального объекта подобным же способом – включением данных о референции в общую структуру метаданных, характеризующих цифровой объект. Кроме того, на основании анализа структуры данных цифровой объект, не содержащий референциальных данных в структуре метаданных, может быть соотнесен с другим цифровым объектом, содержащим такие данные, т.е. может быть установлено, что это различные репрезентации одного и того же объекта. Например, алгоритм распознавания изображений способен установить, что на некоей фотографии изображен тот же самый объект, что и на «эталонной» фотографии, в структуре метаданных которой содержится отсылка к конкретному материальному объекту.

Таким образом, референция цифровых объектов к материальным объектам должна основываться на соответствующих записях в метаданных, постулирующих соответствующую референцию в форме ссылки на реестры материальных объектов и инвентарные списки, обладающих юридической значимостью. Также это могут быть ссылки на другие цифровые объекты, которые, в конечном счете, способны привести нас к правоустанавливающим документам.

Усложнение структуры метаданных, их общей связности, регламентация правил обработки метаданных способны усилить

референциальную связь между объектом и его цифровой копией. Это означает, что чем выше количество атрибутов метаданных, чем сложнее структура метаданных как форм отношения (линков, вложений и т.п.) между различными типами данных, тем сложнее совершить подмену референтного объекта.

Важно понимание, что особенностью цифровых объектов, влияющей на их идентификацию, является возможность получения точной копии объекта без потери качества и каких-либо изменений в исходном объекте. Другими словами, один и тот же цифровой объект может быть представлен неограниченным количеством копий, при этом копии могут быть как полными (как в случае с двумя одинаковыми файлами), так и содержательными (один текст документа, разные форматы), или частичными. В любом случае для таких объектов роль референции при определенном статусе документов в этом контексте является определяющей, подчас единственным технически оптимальным способом, что повышает роль референции и создает предпосылки для спроса на всевозможные реестры и т.д.

Подходы к классификации материальных и цифровых объектов, представляющих культурную ценность

Существует немалое количество классификаций материальных объектов, в том числе обладающих культурной⁶¹ ценностью. В основу таких классификаций кладутся различные критерии: материал изделия, размер, предназначение (эстетическое и/или практическое), стиль и жанр, техника изготовления и т.п. Создание каждой классификации обычно преследует определенную цель – академическую (логически корректная классификация на основании существенных признаков для отображения истинной структуры вещей), пропедевтическую (легкость усвоения материала) или экономическую (удобство учета и инвентаризации).

Полноценная классификация материальных объектов в идеале должна содержать все классы вещей, включая единичные классы, что представляется трудновыполнимой задачей.

61 Термин «культура» в данном значении трактуется в самом широком смысле, включая науку и образование

Подобная классификация, по сути, будет компиляцией всех известных штатных инвентарных категорий. Создание подобной классификации не является необходимой задачей. Достаточно выделить самые общие категории. Существует множество способов деления понятий вещей по разным основаниям. Приведем некоторые из них:

- *Основание способа происхождения материального объекта*
 1. Природные объекты (живые и неживые)
 2. Артефакты
- *Основание цели материального объекта в обществе*
 1. Предметы быта
 2. Предметы эстетические
 3. Предметы композитной природы (бытовые и эстетические в одном)
- *Основание природы вещества изготовления*
 1. Неорганический (минералы и металлы)
 2. Органический (деревья, травы, кости, шкуры, потроха, жилы животных и т.п.)
- *Основание характера темпоральной природы объекта*
 1. Онтологически-условно⁶² статичные объекты (вещи).
 2. Онтологически динамические объекты (волновые колебания)
- *Основание движимости/недвижимости объекта по его назначению.*
 1. Недвижимые объекты
 2. Перемещаемые объекты
 3. Самодвижущиеся объекты (транспортные средства и механизмы)

Обратим внимание на классификацию объектов культурного наследия (памятников истории и культуры), фигурирующую в законодательстве РФ, которые определяются, как объекты недвижимого имущества со связанными с ними произведениями искусства. Согласно Федеральному закону РФ «Об объектах культурного наследия (памятниках истории и культуры) народов Российской Федерации» от 25 июня 2002

⁶² «Условно» потому, что любая вещь подвержена физическим изменениям

года N 73-ФЗ (в ред. Федеральных законов от 27.02.2003 N 29-ФЗ, от 22.08.2004 N 122-ФЗ (ред. 29.12.2004), от 03.06.2005 N 57-ФЗ, от 31.12.2005 N 199-ФЗ), к объектам культурного наследия (памятникам истории и культуры) народов Российской Федерации относятся «объекты недвижимого имущества со связанными с ними произведениями живописи, скульптуры, декоративно-прикладного искусства, объектами науки и техники и иными предметами материальной культуры, возникшие в результате исторических событий, представляющие собой ценность с точки зрения истории, археологии, архитектуры, градостроительства, искусства, науки и техники, эстетики, этнологии или антропологии, социальной культуры и являющиеся свидетельством эпох и цивилизаций, подлинными источниками информации о зарождении и развитии культуры».

Объекты культурного наследия в соответствии с Федеральным законом подразделяются на следующие виды:

памятники – отдельные постройки, здания и сооружения с исторически сложившимися территориями (в том числе памятники религиозного назначения: церкви, колокольни, часовни, костелы, кирхи, мечети, буддистские храмы, пагоды, синагоги, молельные дома и другие объекты, специально предназначенные для богослужений); мемориальные кварталы; мавзолеи, отдельные захоронения; произведения монументального искусства; объекты науки и техники, включая военные; частично или полностью скрытые в земле или под водой следы существования человека, включая все движимые предметы, имеющие к ним отношение, основным или одним из основных источников информации о которых являются археологические раскопки или находки (далее – объекты археологического наследия);

ансамбли – четко локализуемые на исторически сложившихся территориях группы изолированных или объединенных памятников, строений и сооружений фортификационного, дворцового, жилого, общественного, административного, торгового, производственного, научного, учебного назначения, а также

памятников и сооружений религиозного назначения (храмовые комплексы, дацаны, монастыри, подворья), в том числе фрагменты исторических планировок и застроек поселений, которые могут быть отнесены к градостроительным ансамблям; произведения ландшафтной архитектуры и садово-паркового искусства (сады, парки, скверы, бульвары), некрополи;

достопримечательные места – творения, созданные человеком, или совместные творения человека и природы, в том числе места бытования народных художественных промыслов; центры исторических поселений или фрагменты градостроительной планировки и застройки; памятные места, культурные и природные ландшафты, связанные с историей формирования народов и иных этнических общностей на территории Российской Федерации, историческими (в том числе военными) событиями, жизнью выдающихся исторических личностей; культурные слои, остатки построек древних городов, городищ, селищ, стоянок; места совершения религиозных обрядов.

Приведенная классификация имеет правоустанавливающее значение, но оставляет желать лучшего как с точки зрения корректности деления понятий, так и с точки зрения оснований деления понятий. Хотя нельзя исключать, что оно достаточно близко к сложившейся практике отнесения перечисленных видов объектов к той или иной категории.

Использование *прагматических оснований* при создании классификаций обычно вредит научной истине, поскольку прагматические свойства нельзя назвать существенными свойствами объектов самих по себе, но свойствами относительными – поскольку эти свойства есть характеристика взаимодействия материального объекта с окружающей средой. Приведем пример попытки создания подобной классификации: «Пища бывает двоякой – хлеб и прочая еда... Хлеб бывает разных сортов: пресный, квашеный, печеный в золе, поджаренный, пористый, печеный в печи, в виде пирога, сладкого пирожного, а также пшеничный, ячменный, из пшеничной муки мелкого помола, и много других сортов. Прочая еда как бы добавляется к хлебу,

и ее мы можем в общем назвать съестным. Существует много видов съестного. Это мясо, закуски, фрукты, овощи, мед»⁶³.

Можно долго иронизировать над данной классификацией, однако она отражает культурно-бытовые реалии эпохи европейского Средневековья: хлебобулочные изделия действительно являлись основной пищей. Таким образом, *прагматические критерии скорее отображают реальные виды человеческой деятельности*, в которых фигурируют те или иные материальные объекты. Этот критерий является уместным для целей настоящего исследования, поскольку материальные объекты следует представить в соответствии с существующими способами их представления в цифровом виде, что можно интерпретировать как различные виды технической деятельности (методы записи).

Классификация материальных культурных объектов для целей настоящего исследования должна основываться на прагматическом критерии, а именно способе создания цифрового объекта на основании материального объекта. Иными словами, не столь важна корректная классификация материальных объектов, сколько классификация объектов цифровых *по характеру представляемых данных*. Этот характер представляемых данных и будет служить критерием для классификации объектов материального мира. Подобный подход мы можем обнаружить в работе В.И. Ленина, а именно в его знаменитом определении материи: *«Материя есть философская категория для обозначения объективной реальности, которая дана человеку в ощущениях его, которая копируется, фотографируется, отображается нашими ощущениями, существуя независимо от них»*⁶⁴.

В настоящее время существуют всего три способа репрезентации цифровых копий материальных объектов для человеческих органов чувств: визуальный, слуховой и осязательный. Успешно моделировать и представлять вкусовые, обонятельные, проприоцептивные и прочие характеристики наука пока не научилась.

63 Гуго Сен-Викторский. Семь книг назидательного обучения, или Дидаскалион. СПб.: Антология средневековой мысли, 2001. С. 323.

64 Ленин В.И. Материализм и эмпириокритицизм. М.: 1961. С. 131.

Чувственный способ данности	Способы репрезентации цифровых объектов
Визуальный	<ul style="list-style-type: none"> • статичные изображения (картинки, тексты, распечатанные 3-D макеты, и т.п.) • динамические изображения (видеозаписи, тексты в виде бегущих строк и т.п.)
Слуховой	<ul style="list-style-type: none"> • аудиозаписи
Осязательный	<ul style="list-style-type: none"> • аудиозаписи (т.к. ощущаются органами осязания как механические колебания) • текстовые интерфейсы для слепых • интерфейсы игольчато-пиксельных изображений • макеты, распечатанные на 3-D принтере

Предыдущую классификацию по способам чувственной данности следует иметь в виду для понимания категориальной структуры представления материальных объектов в цифровом виде.

Категоризация объектов материального мира по видам человеческой деятельности

Для практических целей предлагается следующая система категорий по основанию вида деятельности или значимости для человеческой культуры.

1 Архитектурные объекты

1.1 Здания и строения

1.2 Транспортные сооружения

1.2.1 мосты

1.2.2 тоннели

1.2.3 галереи

1.2.4 акведуки

1.2.5 лестницы

1.2.6 галереи

1.3 Ландшафтные комплексы (сады, парки)

1.4 Инженерно-промышленные сооружения (вышки, газгольдеры, геодезические купола и т.п.)

1.5 Подземные сооружения (бункеры и катакомбы)

1.6 Архитектурные комплексы зданий

1.7 Архитектурно-ландшафтные комплексы

2 Исторические и природные памятники

2.1 Кенотафы (могильные плиты, склепы, обелиски, керексуры)

2.1.1 надгробия

2.1.2 склепы

2.1.3 обелиски

2.1.4 керексуры (курганы)

2.2 Поминальные камни и комплексы, культовые и исторические надписи и изображения

2.2.1 оленные камни

2.2.2 мегалиты

2.2.3 менгиры

2.2.4 бал-балы

2.2.5 дольмены

2.2.6 петроглифы

2.3 Памятники природы

2.3.1 священные скалы

2.3.2 священные останцы

2.3.4 древние почитаемые деревья

3 Произведения изобразительного искусства

3.1 Живопись

3.1.1 станковая (полотна, создаваемые на мольберте)

3.1.2. монументальная (фрески, мозаика панно, роспись, витраж)

3.1.2.1 фреска

3.1.2.2 мозаика

3.1.2.3 панно

3.1.2.4 роспись

3.1.2.5 витраж

3.1.3 графика

3.2. Скульптура

3.2.1 рельефы

3.2.1.1 горельеф

3.2.1.2 барельеф

3.2.1.3 контррельеф

3.2.2 памятники

- 3.2.3 монументы
- 3.2.4 мемориалы
- 3.3 Граффити
- 3.4 Механические произведения наукоемкого искусства
- 3.5 Механизмы, обладающие культурной ценностью
- 3.6 Холодное оружие, являющееся произведением искусства
- 3.7 Холодное оружие, представляющее культурную и историческую ценность
- 4 Произведения декоративно-прикладного искусства
 - 4.1 Резьба (по кости, дереву)
 - 4.2 Чеканка по металлу
 - 4.3 Кружева
 - 4.4 Папье-маше
 - 4.5 Художественно обработанный камень
 - 4.6 Декорированная береста
 - 4.7 Ткацкие произведения (ткань, ковры)
 - 4.8 Инкрустация (камень, дерево, кость)
 - 4.9 Раковины моллюсков обработанные
 - 4.10 Литье
 - 4.11 Скобяные изделия
 - 4.12 Кузнечные произведения
- 5 Письменные и печатные тексты
 - 5.1 Книги
 - 5.2 Статьи в периодических изданиях и сборниках
 - 5.3 Справочники
 - 5.4 Словари
 - 5.5 Брошюры
 - 5.6 Партитуры
 - 5.7 Письма
 - 5.8 Документы
 - 5.9 Личные дневники, блокноты и тетради
- 6. Научные и инженерные рукотворные и печатные графические изображения
 - 6.1 Чертежи
 - 6.2 Схемы
 - 6.3 Планы

- 6.4 Графики
- 6.5 Диаграммы
- 6.6 Таблицы
- 6.7 Карты
- 6.8 Программы для ЭВМ
- 7. Материальные отображения форм стационарных объектов
 - 7.1 Фотографии
 - 7.2 Изображение на первичном фотографическом носителе (проявленные и непроявленные)
 - 7.3 Дагерротипы
 - 7.4 Киноплёнка
 - 7.5 Отображение на первичном магнитном носителе для аналоговых записей фото и видео
 - 7.6 Слепки
 - 7.7 Оттиски
 - 7.8 Светокопии (фотокопии)
 - 7.9 Масштабные макеты
 - 7.10 Масштабные модели
- 8. Аналоговые аудиозаписи
 - 8.1 Виниловые пластинки
 - 8.2 Парафиновые барабаны
 - 8.3 Барабаны музыкальных шкатулок
 - 8.4 Магнитная плёнка
 - 8.5 Магнитная проволока

Категоризация цифровых объектов по форматам записи

- 1. 3-D файлы
 - 3DM (Rhinceros)
 - 3DS (3D Shape)
 - A3D (Компас-3D (файл сборки))
 - ASM (Pro/Engineer)
 - BLEND (Blender)
 - C4D (Cinema 4D)
 - DDS (DirectDraw Surface)
 - DGN (Bentley Systems MicroStation)
 - DWG (AutoCAD)

- DXF (AutoCAD)
- FBX (Alias)
- LWO (LightWave Object File)
- LWS (LightWave Scene File)
- M3D (Компас-3D (файл модели))
- MA (Maya)
- MAX (3D Studio Max)
- MB (Maya)
- MOVIE.BYU
- Obj (WaveFront Object)
- PZ3 (Pozer)
- PRT (Pro/Engineer)
- RIB (RenderMan)
- SLDASM (SolidWorks файл сборки)
- SLDPRT (SolidWorks файл модели)
- STP (ISO 10303-21)
- U3D (3D Industry Forum универсальный формат файла трехмерной графики, использующий сжатие данных)
- VRML (стандартизированный формат файлов для демонстрации трехмерной интерактивной векторной графики)
- VUE (Vue) (стандарт для 3D моделей)
- WIRE (Alias STUDIO Tools)(стандарт для общего описания информации на уровне приложений)
- X (DirectX) (набор API⁶⁵, разработанных для решения задач, связанных с программированием под Microsoft Windows.)
- X3D (стандарт ISO, предназначенный для работы с трехмерной графикой в реальном времени.)
- Z3d (Zanoza Modeler) (3D моделирование приложений)
- MDL (стандарт для создания сайтов)
- NIF (NetImmerse/Gamebryo file) (3D-модель, используемая в игровом движке)

65 Набор готовых классов, процедур, функций, структур и констант, предоставляемых приложением (библиотекой, сервисом) для использования во внешних программных продуктах. Используется программистами при написании всевозможных приложений.

- Md (TrumGame ModelEditor) (формат для изображения, используемый для 3D проектирования)

- SPR (SCAD) (SPR – расширение файлов для программы SCAD (Structure CAD): приложение которое использует файл .SPR, формат файла используется во многих играх)

- UNITY (Unity 3D) (используется для разработки двух- и трехмерных приложений и игр, работающих под операционными системами Windows, OS X. Созданные с помощью Unity приложения работают под операционными системами Windows, OS X, Windows Phone, Android, Apple iOS, Linux)

2. Фонограммы (звукозаписи)

- AA – Audible Audio Book File

- AAC – Advanced Audio Codin (формат аудиофайла и алгоритма кодирования с меньшей, чем у MP3, потерей качества при кодировании с одинаковыми битрейтами)

- AMR – Adaptive Multi-Rate (стандарт кодирования звуковых файлов, специально предназначенный для сжатия сигнала в речевом диапазоне частот)

- APE – Monkey’s Audio (популярный формат кодирования цифрового звука без потерь. Кодек распространяется в виде бесплатного одноименного программного обеспечения для Microsoft Windows, а также плагинов к популярным медиаплеерам)

- ASF – Advanced Systems Format (разработанный фирмой Microsoft формат файлов, содержащих потоковое аудио и видео)

- CDA – CD Digital Audio

- FLAC – Free Lossless Audio Codec (популярный свободный кодек, предназначенный для сжатия аудиоданных без потерь.)

- LAV – Liquid Audio Voucher

- MP3 (кодек третьего уровня, разработанный командой MPEG, лицензируемый формат файла для хранения аудиоинформации)

- MT9 – Music 2.0 (формат, разработанный корейскими специалистами . Также имеет название Music 2.0. MT9 позволяет слушателям регулировать громкость для каждого канала – такого как гитара, барабаны, бас и вокал – заглушка или усиление

их любимых частей. В дальнейшем ожидается поддержка МТ9 сотовыми телефонам)

- Ogg (открытый стандарт формата мультимедиаконтейнера, являющийся основным файловым и потоковым форматом для мультимедиакодексов фонда Xiph)

- VOC – Creative SoundBlaster (формат, используемый для диктофонных записей)

- WAV – Windows PCM/ADPCM (формат файла-контейнера для хранения записи оцифрованного аудиопотока)

- WMA – Windows Media Audio (лицензируемый формат файла, разработанный компанией Microsoft для хранения и трансляции аудиоинформации)

3. Синтезируемые фонограммы

- MIDI (цифровой интерфейс музыкальных инструментов) – стандарт цифровой звукозаписи на формат обмена данными между электронными музыкальными инструментами

- MOD (формат трекерной музыки)

- STM, S3M (Stream Tracker) (формат музыкальных файлов, использовавшийся трекером ScreamTracker 3 – один из трех наиболее распространенных форматов трекерной музыки. S3M является развитием формата STM)

- MMF (это мультимедийный формат данных, разработанный компанией Yamaha для эффективного и компактного хранения мультимедийных данных)

- Gtr, Gr3, Gr4, Gr5, Grx (Guitar Pro, Tux Guitar) (форматы для гитарных партитур и нотных табулатур)

4. Растровые изображения

- BMP (Windows or OS/2 bitmap) (формат хранения растровых изображений, разработанный компанией Microsoft. Файлы формата BMP могут иметь расширения .bmp, .dib и .rle)

- CPT (Corel Photo-Paint)

- GIF (Graphics Interchange Format) (формат для движущегося изображения)

- HDR (High Dynamic Range) (технологии работы с изображениями и видео, диапазон яркости которых превышает возможности стандартных технологий)

- JPEG, JPG, JPE (Joint Photographic Experts Group) (формат Windows для хранения и передачи изображений)

- JPEG 2000 (.jp2) (графический формат, который вместо дискретного косинусного преобразования, применяемого в формате JPEG, использует технологию вейвлет-преобразования, основывающуюся на представлении сигнала в виде суперпозиции базовых функций – волновых пакетов)

- PCX (ZSoft PaintBrush) (стандарт представления графической информации, разработанный компанией ZSoft Corporation)

- PDF (Adobe) (пакет программ, выпускаемый компанией Adobe Systems для создания и просмотра электронных публикаций в формате PDF)

- PDN (Paint.NET Image) (бесплатный растровый графический редактор для Windows NT, основанный на .NET Framework)

- PNG (Portable Network Graphics) (растровый формат хранения графической информации, использующий сжатие без потерь по алгоритму Deflate.)

- PSD (Photoshop document) (растровый формат хранения графической информации, использующий сжатие без потерь, созданный специально для программы Adobe Photoshop и поддерживающий все его возможности)

- Raw (формат данных, содержащий необработанные (или обработанные в минимальной степени) данные, что позволяет избежать потерь информации, и не имеющий четкой спецификации)

- TGA (Truevision Targa) (.TGA, .tpic) (растровый графический формат; первоначально был создан компанией Truevision Inc. для графических адаптеров собственного производства еще в 1984, в дальнейшем стал популярен на самых разных платформах, особенно в области обработки видео, анимации)

- TIFF, TIF (Tagged Image Format) (формат хранения растровых графических изображений. TIFF стал популярным форматом для хранения изображений с большой глубиной цвета)

- WDP, HDP (Windows Media Photo) (формат кодирования и файловый формат для фотографий, ранее разработанный и запатентованный корпорацией Microsoft)

- XPM (X pixmap) (текстовый формат графических файлов, используется серверной стороной системы X Window)

5. Векторные изображения

- AI (Adobe Illustrator) (формат для разнообразной векторной графики)

- CDR (Corel Draw) (файл проекта, созданный в программе CorelDRAW, который содержит векторное изображение или растровый рисунок. Данный формат файла разработан компанией Corel для использования в собственных программных продуктах)

- CGM (Computer Graphics Metafile) (открытый формат и международный стандарт для хранения и обмена графическими данными)

- EPS (Encapsulated PostScript format) (формат файлов, базирующийся на подмножестве языка PostScript)

- PDF (Portable Document Format)

- PS (PostScript) (формат для описания страниц, в основном используемый в настольных издательских системах)

- SVG (Scalable Vector Graphics) (предназначен для описания двумерной векторной и смешанной векторно/растровой графики в формате XML)

- SAI (Paint Tool SAI) (используется при работе с анимацией)

- WMF (Windows Metafile) (мультимедийный набор от Microsoft для создания и распространения аудио и видеофайлов для Windows. Включает набор средств разработки и интерфейса программирования приложений)

- EMF (Extended Metafile)⁶⁶ (универсальный формат векторных графических файлов для Windows приложений. Используется для хранения коллекции графических изображений Microsoft Clip Gallery)

6. Анимация

- ANI (Windows Animated Cursor) (графический файл для курсоров мыши)

⁶⁶ Расширение файлов, типы файлов URL: <http://fileext.ru/> (дата обращения: 13.12.2015).

- APNG (Animated PNG) (формат изображений, основанный на формате PNG и предусматривающий возможность хранения анимации, аналогичной используемой в формате GIF)

- FLC (Autodesk Animator Pro) (формат для 2D анимации)
- FLI (Autodesk Animator) (формат для 2D анимации)
- GIF

- MNG (Multiple-image Network Graphics) (формат графических файлов. Поддерживает все возможности алгоритмов сжатия PNG и JPEG (в том числе полупрозрачность и гамма-коррекцию) для создания анимированных изображений. Поддерживается в браузере Konqueror, поддерживался в браузере Mozilla 2000 до 2003 года)

- SMIL, презентации (открытый аналог Flash) (язык разметки для создания интерактивных мультимедийных презентаций)

- SVG (Scalable Vector Graphics) – презентационный формат, ориентированный на векторную графику

- SWF (Adobe Flash, векторная графика)

- WLMP (Windows Movie Maker)⁶⁷ (формат для работы с анимацией от Windows)

7. Видео

- ASF

- AVI (Audio Video Interleave) (RIFF-медиаконтейнер, впервые использованный Microsoft в 1992 году в пакете Video for Windows.)

- HDrip (расшифровывается HDTV как High Definition TeleVision (телевидение высокой четкости), и его самая главная отличительная черта – это исключительно высокое разрешение: 1280*720 (стандарт 720p) или 1920*1080 (стандарт 1080i))

- BIK (BinkVideo) (проприетарный формат файла, медиаконтейнер, разработанный американской компанией RAD Game Tools для использования преимущественно в компьютерных играх для разных игровых платформ)

- FLV (Flash Video) (формат файлов, медиаконтейнер, используемый для передачи видео через интернет)

67 FileWiki URL: <http://filewiki.ru/> (дата обращения: 13.12.2015).

- Matroska (.mkv) (расширения файлов Matroska: .mkv – для видео (с субтитрами и звуком), .mka – для аудиофайлов, .mks – для субтитров и .mk3d – для 3D-видео. Возможности формата, закладываемые в Matroska: трансляция через интернет (протоколы HTTP и RTP); быстрая перемотка по файлу; устойчивость к ошибкам)
- MPEG-TS (является протоколом для передачи аудио- и видеоданных, описанным в MPEG2)
 - MXF (Material eXchange Format) (формат обмена данными – медиаконтейнер для профессионального хранения и обработки видео)
 - Ogg (для Tarkin и Theora)
 - QuickTime (.mov, .qt) (проприетарная технология компании Apple, разработанная в 1991 году для воспроизведения цифрового видео)
 - RealMedia (проприетарный стандарт на потоковое вещание и на формат медиафайлов, принадлежащий фирме «RealNetworks Products»)
 - RIFF-AVI (один из форматов файлов-контейнеров для хранения потоковых мультимедиа-данных (видео, аудио, возможно текст). Наиболее известными форматами, использующими RIFF в качестве контейнера, являются: AVI (видео), WAV (аудио), RMI (MIDI-треки))
 - SWF (ShockWave File) (проприетарный формат для флеш-анимации, векторной графики, видео и аудио в сети Интернет. Картинка, сохраненная в этом формате, масштабируется без видимых искажений, видеоролик имеет небольшой размер, происходит более быстрая загрузка видеофайла и его воспроизведение)
 - VOB (DVD-Video File) (формат файлов, используемый для хранения DVD-Video. VOB – это мультимедийный контейнерный формат файла, основанный на MPEG-2, и способный содержать в себе несколько потоков видео/аудио, субтитры, а также меню фильма)
 - WMV (название системы видеокодирования, разработанной компанией Microsoft для хранения и трансляции видеoinформации)

- AVCHD (.mts)⁶⁸ (формат записи видео, предусматривающий запись на безленточные носители, основанный на кодеке MPEG-4 AVC/H264)

8. Графические растровые форматы

- TIFF (один из основных форматов для хранения фотографий. Он используется в типографской продукции, при сканировании и отправке факсов)

- JPEG

- DjVu (технология сжатия изображений с потерями, разработанная специально для хранения сканированных документов – книг, журналов, рукописей и прочее, где обилие формул, схем, рисунков и рукописных символов делает чрезвычайно трудоемким их полноценное распознавание)

9. Представление текстов с оформлением

- RTF (проприетарный межплатформенный формат хранения размеченных текстовых документов, предложенный группой программистов, основавшими компании Microsoft и Adobe)

- OPF FlipBook (открытый формат электронных книг – стандартная версия программы FlipAlbum создает альбомы-книги FlipBooks с расширением OPF)

- FB2 FB3 TEX (форматы представления электронных версий книг в виде XML-документов, где каждый элемент книги описывается своими тегами; стандарт призван обеспечить совместимость с любыми устройствами и форматами)

- PDF LIT PostScript и EPS (формат файлов, базирующийся на подмножестве языка PostScript)

- ExeBook (позволяет хранить книги в форме самоисполняемого файла)

- PRC (формат PRC представляет собой формат электронной книги, созданный компанией Mobipocket и используемый исключительно для устройств Palm, которые поддерживают форматы PRC или PDB. Файлы PRC также можно просмотреть на компьютере с помощью программы Mobipocket Reader Desktop)

68 <http://dotwhat.net/>

• ePUB (открытый формат электронных версий книг, разработанный Международным форумом по цифровым публикациям)

10. Простой текст

• TXT (компьютерный файл, содержащий текстовые данные)

Отношения между материальными объектами, представляющими культурную ценность, и типами цифровых форматов

(В таблице сделана попытка сопоставления перечня видов деятельности со способами их репрезентации в цифровой среде.)

<p>1. Архитектурные объекты</p> <p>1.1 Здания и строения</p> <p>1.2 Транспортные сооружения</p> <p>1.2.1 мосты</p> <p>1.2.2 тоннели</p> <p>1.2.3 галереи</p> <p>1.2.4 акведуки</p> <p>1.2.5 лестницы</p> <p>1.2.6 галереи</p> <p>1.3 Ландшафтные комплексы (сады, парки)</p> <p>1.4 Инженерно-промышленные сооружения (вышки, газгольдеры, геодезические купола и т.п.)</p> <p>1.5 Подземные сооружения (бункеры и катакомбы)</p> <p>1.6 Архитектурные комплексы зданий</p> <p>1.7 Архитектурно-ландшафтные комплексы</p>	<p>1. 3-D модели</p> <ul style="list-style-type: none"> • 3DM (Rhinceros) • 3DS (3D Shape) • A3D (Компас-3D (файл сборки)) • ASM (Pro/Engineer) • BLEND (Blender) • C4D (Cinema 4D) • DDS (DirectDraw Surface)⁶⁹ • DGN (Bentley Systems MicroStation) • DWG (AutoCAD) • DXF (AutoCAD) • FBX (Alias) • LWO (LightWave Object File) • LWS (LightWave Scene File) • M3D (Компас-3D (файл модели)) • MA (Maya) • MAX (3D Studio Max) • MB (Maya) • MOVIE.BYU • Obj (WaveFront Object) • PZ3 (Pozer) • PRT (Pro/Engineer) • RIB (RenderMan) • SLDASM (SolidWorks файл сборки) • SLDPRT (SolidWorks файл модели) • STP (ISO 10303-21) • U3D • VRML • VUE (Vue)
---	---

69 dotWhat? URL: <http://dotwhat.net/> (дата обращения: 13.12.2015)

	<ul style="list-style-type: none"> • WIRE (Alias STUDIO Tools) • X (DirectX) • X3D • Z3d (Zanoza Modeler) • MDL • NIF (NetImmerse/Gamebryo file) • Md (TrumGame ModelEditor) • SPR (SCAD) • UNITY (Unity 3D) <p>2. Растровые изображения</p> <ul style="list-style-type: none"> • BMP (Windows or OS/2 bitmap) • CPT (Corel Photo-Paint) • GIF (Graphics Interchange Format) • HDR (High Dynamic Range) • JPEG, JPG, JPE (Joint Photographic Experts Group) • JPEG 2000 (.jp2) • PCX (ZSoft PaintBrush) • PDF (Adobe) • PDN (Paint.NET Image) • PNG (Portable Network Graphics) • PSD (Photoshop document) • Raw • TGA (Truevision Targa) (.TGA, .tpic) • TIFF, TIF (Tagged Image Format) • WDP, HDP (Windows Media Photo) • XPM (X pixmap) <p>3. Видео</p> <ul style="list-style-type: none"> • QuickTime (.mov, .qt) • RealMedia • RIFF-AVI • Smacker (*.smk) • SWF (ShockWave File) • VOB (DVD-Video File) • WMV • AVCHD (.mts)
<p>2 Исторические и природные памятники</p> <p>2.1 Кенотафы (могильные плиты, склепы, обелиски, керексуры)</p>	<p>1. 3-D модели</p> <p>2. Растровые изображения</p> <p>3. Видео</p>

<p>2.1.1 надгорбия 2.1.2 склепы 2.1.3 обелиски 2.1.4 керексуры (курганы) 2.2 Поминальные камни и комплексы, культовые и исторические надписи и изображения 2.2.1 оленные камни 2.2.2 мегалиты 2.2.3 менгиры 2.2.4 бал-балы 2.2.5 дольмены 2.2.6 петроглифы 2.3 Памятники природы 2.3.1 священные скалы 2.3.2 священные останцы 2.3.4 древние почитае- мые деревья 3 Произведения изобрази- тельного искусства 3.1 Живопись 3.1.1 станковая (по- лотна, создаваемые на мольберте) 3.1.2 монументальная (фрески, мозаика панно, ро- спись, витраж) 3.1.2.1 фреска 3.1.2.2 мозаика 3.1.2.3 панно 3.1.2.4 роспись 3.1.2.5 витраж 3.1.3 графика 3.2. Скульптура 3.2.1 рельефы 3.2.1.1 гроельеф 3.2.1.2 барельеф 3.2.1.3 контррельеф</p>	
--	--

<p>3.2.2 памятники 3.2.3 монументы 3.2.4 мемориалы 3.3 Граффити 3.4 Механические произведения наукоёмкого искусства. 3.5 Механизмы, обладающие культурной ценностью. 3.6 Холодное оружие, являющееся произведением искусства 3.7 Холодное оружие, представляющее культурную и историческую ценность</p>	
<p>4 Произведения декоративно-прикладного искусства 4.1 Резьба (по кости, дереву) 4.2 Чеканка по металлу 4.3 Кружева 4.4 Папье-маше 4.5 Художественно-обработанный камень 4.6 Декорированная береста 4.7 Ткацкие произведения (ткань, ковры) 4.8 Инкрустация (камень, дерево, кость) 4.9 Раковины моллюсков обработанные 4.10 Литье 4.11 Скобяные изделия 4.12 Кузнечные произведения</p>	<p>1. 3-D модели 2. Растровые изображения 3. Видео</p>
<p>5 Письменные и печатные тексты 5.1 Книги 5.2 Статьи в периодических изданиях и сборниках</p>	<p>1. Графические растровые форматы • TIFF • JPEG • DjVu</p>

<p>5.3 Справочники 5.4 Словари 5.5 Брошюры 5.6 Партитуры 5.7 Письма 5.8 Документы 5.9 Личные дневники, блокноты и тетради</p>	<p>2. Графические векторные форматы с оформлением</p> <ul style="list-style-type: none"> • RTF • OPF FlipBook • CHM • SGML • FB2 • FB3 • TEX • PDF • LIT • PostScript и EPS • ExeBook • RBS • PRC • ePUB <p>3. Простой текст</p> <ul style="list-style-type: none"> • TXT <p>4. Аудиофайлы</p>
<p>6 Научные и инженерные рукотворные и печатные графические изображения</p> <p>6.1 Чертежи 6.2 Схемы 6.3 Планы 6.4 Графики 6.5 Диаграммы 6.6 Таблицы 6.7 Карты 6.8 Программы для ЭВМ</p>	<p>1. Графические растровые форматы 2. Графические векторные форматы с оформлением</p>
<p>7 Материальные отображения форм стационарных объектов</p> <p>7.1 Фотографии 7.2 Изображение на первичном фотографическом носителе (проявленные и непроявленные) 7.3 Дагерротипы 7.4 Кинопленка</p>	<p>1. Растровые изображения 2. Графические растровые форматы 3. Графические векторные форматы с оформлением 4. Видео</p>

<p>7.5 Отображение на первичном магнитном носителе для аналоговых записей фото и видео</p> <p>7.6 Слепки</p> <p>7.7 Оттиски</p> <p>7.8 Светокопии (фотокопии)</p> <p>7.9 Масштабные макеты</p> <p>7.10 Масштабные модели</p>	
<p>8. Аналоговые аудиозаписи</p> <p>8.1 Виниловые пластинки</p> <p>8.2 Парафиновые барабаны</p> <p>8.3 Барабаны музыкальных шкатулок</p> <p>8.4 Магнитная пленка</p> <p>8.5 Магнитная проволока</p>	<p>Фонограммы (звукозаписи)</p> <ul style="list-style-type: none"> • AA – Audible Audio Book File • AAC – Advanced Audio Coding • AMR – Adaptive Multi-Rate • APE – Monkey’s Audio • ASF • CDA – CD Digital Audio • FLAC – fla • LAV – Liquid Audio Voucher • MP3 • MT9 – Music 2.0 • Ogg • VOC – Creative SoundBlaster • WAV – Windows PCM/ADPCM • WMA

Юридические особенности

Возможность конвертирования различных форматов записи влечет нечеткость границ между цифровыми объектами. На примере оцифровки графиков/чертежей/схем можно увидеть, что изображение как объект материального мира и цифровой объект в формате изображения, который позволяет демонстрировать его в сети, суть разные объекты. Цифровой объект в формате изображения представляется продуктом актуальным только тогда, когда пользователю необходимо поделиться конвертируемой в этот формат информацией. Пользователь, ко всему прочему, обладает свободой выбора формата, в зависимости от того, как ему удобнее делиться информацией с другими пользователями сети. Необходимо отметить, что аудиофайлы могут воспроизводить тексты художественного и информационного содержания.

Из приведенной выше классификации можно сделать вывод, что объекты материального мира могут быть отображены и выложены в сеть не только при помощи напрямую соответствующего им формата (текст – текстовый файл), но и в форматах объектов, выражение в которых возможно, если пользователь сочтет нужным (и более удобным) преподнести информацию *в нетипичном для конкретного объекта в сложившихся способах социокультурной репрезентации информации формате* (аудиофайл, отображающий текстовый документ). Таким образом, очевидно наличие вариативности представленности исходного объекта в цифровой форме.

Несмотря на наличие существенного различия между исходным объектом и его оцифрованным воплощением, оцифрованный вариант обладает уникальным содержанием, которое передается в различных форматах. Чрезвычайно важно понимать, что законодательство об авторском и интеллектуальном праве охраняет те результаты интеллектуальной деятельности, которые были созданы при помощи интеллектуальных и творческих усилий человека, который является их автором. Согласно ст. 1228 ГК РФ, «автором результата интеллектуальной деятельности признается гражданин, творческим трудом которого создан такой результат. Не признаются авторами результата интеллектуальной деятельности граждане, не внесшие личного творческого вклада в создание такого результата, в том числе оказавшие его автору только техническое, консультационное, организационное или материальное содействие или помощь либо только способствовавшие оформлению прав на такой результат или его использованию, а также граждане, осуществлявшие контроль за выполнением соответствующих работ».

Согласно ст. 1225. ГК РФ «Охраняемые результаты интеллектуальной деятельности и средства индивидуализации»:

1. Результатами интеллектуальной деятельности и приравненными к ним средствами индивидуализации юридических лиц, товаров, работ, услуг и предприятий, которым предоставляется правовая охрана (интеллектуальной собственностью), являются:

- 1) произведения науки, литературы и искусства;
- 2) программы для электронных вычислительных машин (программы для ЭВМ);
- 3) базы данных;
- 4) исполнения;
- 5) фонограммы;
- 6) сообщение в эфир или по кабелю радио- или телепередач (вещание организаций эфирного или кабельного вещания);
- 7) изобретения;
- 8) полезные модели;
- 9) промышленные образцы;
- 10) селекционные достижения;
- 11) топологии интегральных микросхем;
- 12) секреты производства (ноу-хау);
- 13) фирменные наименования;
- 14) товарные знаки и знаки обслуживания;
- 15) наименования мест происхождения товаров;
- 16) коммерческие обозначения.

Согласно Бернской конвенции, понятие «литературные и художественные произведения» «охватывает любую продукцию в области литературы, науки и искусства, вне зависимости от способа и формы ее выражения, включая: книги, брошюры и другие письменные произведения; лекции, обращения, проповеди и другие подобного рода произведения; драматические и музыкально-драматические произведения; хореографические произведения и пантомимы; музыкальные сочинения с текстом или без текста; кинематографические произведения, к которым приравниваются произведения, выраженные способом, аналогичным кинематографии; рисунки, произведения живописи, архитектуры, скульптуры, гравирования и литографии; фотографические произведения, к которым приравниваются произведения, выраженные способом, аналогичным фотографии; произведения прикладного искусства; иллюстрации, карты, планы, эскизы и пластические произведения, относящиеся к географии, топографии, архитектуре или наукам». Таким образом, опираясь на ратифицированную в РФ конвенцию, можно

сделать вывод, что защите результат интеллектуального труда подлежит вне зависимости от формы своего выражения. В период разработки и опубликования Бернской конвенции не существовало возможности оцифровывать объекты материального мира, но оцифровывание в определенный формат есть не что иное, как смена способа выражения материального объекта⁷⁰.

Международное регулирование защиты авторских прав в сфере произведений науки, техники, товарных знаков, то есть объектов, являющихся результатом умственного труда и поддающихся оцифровке, осуществляется при помощи норм «Всемирной декларации по интеллектуальной собственности» от 2000 года⁷¹.

В Декларации термин «интеллектуальная собственность» означает «любую собственность, признаваемую по общему согласию в качестве интеллектуальной по характеру и заслуживающей охраны, включая, но не ограничиваясь научными и техническими изобретениями, литературными или художественными произведениями, товарными знаками и указателями деловых предприятий, промышленными образцами и географическими указаниями»⁷². Таким образом, под охрану норм гражданского законодательства подпадают не только результаты интеллектуальной деятельности в области культуры, но и научно-технические достижения. Следовательно, объекты, отображающие результаты научно-технического развития, также подлежат охране нормами об авторском праве.

Оцифровка объектов представляется единственным на данный момент способом радикального облегчения доступа широкой общественности к культурному содержанию. Конвертирование

70 Всемирная организация интеллектуальной собственности URL: http://www.wipo.int/wipolex/ru/treaties/text.jsp?file_id=283702#a2 (дата обращения: 13.12.2015)

71 Федеральная служба по интеллектуальной собственности (Роспатент) URL: http://www.rupto.ru/docs/interdocs/worldwide_declaration (дата обращения: 13.12.2015)

72 Федеральная служба по интеллектуальной собственности (Роспатент) URL: http://www.rupto.ru/docs/interdocs/worldwide_declaration (дата обращения: 13.12.2015)

цифровых объектов может осуществляться в различных форматах. Формат – спецификация данных, записанных в компьютерном файле, то есть в таком файле, с которым компьютер (вычислительная машина) имеет возможность работать⁷³. Форматирование осуществляется в зависимости от того, какой тип программ может работать с заданным изначально форматом. Разнообразие программ, способных отображать и воспроизводить оцифрованные объекты, обусловлено существованием программного рынка, который в условиях постоянной конкуренции предоставляет пользователю альтернативу в выборе используемого программного обеспечения. Существует систематизированный список расширений (форматов файлов), которые используются при оцифровке объектов.

В ситуации, когда один тип программ представляется отправителю файла не совсем удобным или надежным, файл конвертируется в необходимый формат. Пользователь имеет возможность выбора формата своего файла для дальнейшей загрузки в сеть.

Российское законодательство, несмотря на свое несовершенство в этой сфере, в самом общем виде предусматривает право на защиту результатов интеллектуальной деятельности всеми доступными и не противоречащими закону способами. Согласно ст. 1299 ГК РФ, «техническими средствами защиты авторских прав признаются любые технологии, технические устройства или их компоненты, контролирующие доступ к производству, предотвращающие либо ограничивающие осуществление действий, которые не разрешены автором или иным правообладателем в отношении произведения. В отношении произведений не допускается: 1) осуществление без разрешения автора или иного правообладателя действий, направленных на то, чтобы устранить ограничения использования произведения, установленные путем применения технических средств защиты авторских прав; 2) изготовление, распространение, сдача в прокат, предоставление во временное безвозмездное пользование, импорт, реклама любой технологии, любого технического устройства или их компонентов,

73 URL: <http://www.file-extensions.org/> (дата обращения: 13.12.2015)

использование таких технических средств в целях получения прибыли либо оказание соответствующих услуг, если в результате таких действий становится невозможным использование технических средств защиты авторских прав либо эти технические средства не смогут обеспечить надлежащую защиту указанных прав»⁷⁴.

Таким образом, законодатель предусматривает возможность использования любых доступных и законных технологий для защиты авторского права. Подобного рода формулировка «технологии» была предусмотрена законодателем с учетом развития информационных технологий и сложности контроля за тем, чтобы выложенный в сеть контент использовался пользователями в установленных законом рамках.

Определение основных реестров произведений культуры, науки и образования (ISBN, DOI и др.)

Предпосылки создания новых систем идентификаторов объектов, являющихся результатами интеллектуального труда

В связи с массовым распространением интеллектуальной собственности в сети Интернет стали разрабатываться системы идентификаторов, направленные на хранение и систематизацию объектов интеллектуального труда. Присвоение идентификаторов, направленных на идентификацию (процесс присвоения как субъектам, так и объектам коммуникации определенных уникальных идентификаторов и их сравнение с перечнем присвоенных идентификаторов) размещенного в глобальной сети контента является одним из основных методов упрощения поиска и доступа пользователей к необходимой информации. Важная задача, стоящая перед идентификацией пользовательского контента – фиксация выложенного в сеть материала с целью отслеживания динамики его копирования, представления в других

⁷⁴ КонсультантПлюс URL: <http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=182704> (дата обращения: 13.12.2015)

форматах и контроля за соблюдением законодательства о защите авторских и (или) смежных прав в сети Интернет.

Используемые идентификаторы представляется целесообразным поделить на два типа. Это идентификаторы, специализированные по типовому компоненту репрезентируемой информации или носящие универсальный характер. В России кроме международных реестров существуют национальные реестры, перечень которых указан на сайте Роспатента⁷⁵.

В зависимости от целей, преследуемых правообладателем, оцифрованный объект может получить идентификационные метки в реестре как универсального, так и узкоспециализированного характера.

Идентификация в сети происходит на основании присвоения сформированных идентификаторов каждому из опубликованных в сети объектов. Метаданные как феномен – неотъемлемая часть объекта, с помощью которой происходит идентификация в сети. Идентификация и отображение объектов в так называемых репозиториях необходима для отслеживания динамики изменения объектов в других форматах.

Метаданные об объекте могут быть подвержены изменению: коррекции, оптимизации и конвертации из одного формата записи данных в другой в то время, как сам цифровой объект остается неизменным. Например, после проведения юридически значимого разбирательства на предмет установления автора произведения, авторство может быть приписано другому лицу, которое ранее не указывалось в реестре данных об объекте в качестве автора. В данном примере мы сталкиваемся со случаем изменения фрагмента метаданных о цифровом объекте. Таким образом, при замене автора произведения изменяются метаданные, но не сам цифровой объект. Подобное уточнение позволяет более правильно идентифицировать объект в системе поиска.

Данные и метаданные являются схожими по своему существу феноменами. В зависимости от типа объекта, одно и тоже изменение может быть расценено в структуре данных объектов

⁷⁵ Федеральная служба по интеллектуальной собственности (Роспатент) URL: <http://www.rupto.ru/about/fgis> (дата обращения: 13.12.2015)

различным образом. Например, изменение разрешения изображения в формате JPG влечет за собой изменение метаданных, но не цифрового объекта. Подобного рода последовательные изменения являются динамической характеристикой существования результата интеллектуального труда в сети, что, в принципе, может быть отслежено и учтено. Изменение метаданных в приведенном примере продиктовано тем, что происходит изменение формальных характеристик цифрового объекта, сама сущность объекта не подвергается изменению. Метаданные подвергаются международной стандартизации с целью более быстрого поиска, автоматической обработки, понимания на разных языках, а также максимальной точности описываемого объекта с наименьшей затратой используемых символов. Метаданные выражают уникальность демонстрируемого в сети объекта. Таким образом, можно полагать, что использование метаданных существенно упрощает и автоматизирует процедуру идентификации и формирования систематизированного перечня цифровых объектов и позволяет преодолеть хаос структуры данных о цифровых объектах в глобальной сети.

Стандартизация метаданных зависит от спектра охватываемых стандартом метаданных объектов. Для целей описания библиографических ресурсов созданы форматы семейства MARC и UNIMARC. Последний, в свою очередь, подразделяется на национальные расширения этого формата (так, в России используется RUSMARC); в силу большой сложности форматов семейства MARC для решения задач интеграции данных был разработан формат MODS⁷⁶.

Способ форматирования метаданных зависит от описываемого объекта. Для определенного типа объектов (книга, периодическое издание и т.д.) существуют специальные форматы. Формат может охватывать несколько типов объектов или быть предназначенным для одного конкретного типа. Например,

76 Воройский Ф. С. Информатика. Новый систематизированный словарь-справочник (Вводный курс по информатике и вычислительной технике в терминах). 2-е изд, перераб. и доп. изд. М.: Издательство Либерия, 2001. С. 536

формат MARC позволяет учитывать объекты разных типов, он содержит поля для описания названия, автора, тематики и огромного множества других характеристик (формат MARC позволяет описать сотни характеристик)⁷⁷.

Создание хранилищ оцифрованных копий (отпечатков) объектов материального мира продиктовано несколькими причинами:

1. Распространение, хранение, расширения доступа к результатам интеллектуальной собственности.

2. Охрана авторского права.

3. Борьба с «пиратством».

Однако, в связи с реализацией принципа свободного распространения информации, реализованного в сети Интернет, возникает серьезная проблема, с которой идентификация и привязка оцифрованных объектов к хранилищу данных не справляется.

Реестры могут оказывать лишь косвенную поддержку в борьбе с незаконным использованием результатов интеллектуальной собственности, а именно определять, за каким юридическим или физическим лицом закреплен пакет авторских прав. С опорой на национальные и международные нормативно-правовые акты, методом анализа и соотнесения субъектов, осуществляющих авторские полномочия, можно определить, имеет ли место в каждом конкретном случае нарушение авторского права, или же факт правонарушения не состоялся. Необходимо понимать, что только одно создание реестров *не позволит пресечь пиратские атаки на результаты интеллектуальной собственности*. Концепция непрерывного и постоянного отслеживания динамики представленных объектов адекватна лишь в случаях, когда необходимо разрешить спор об авторстве или отследить тенденции спроса пользователей на различный контент.

Реестр осуществляет функцию закрепления репрезентации (отпечатка) объекта за самим материальным объектом. Отображение оцифрованной копии объекта материального

77 MARC STANDARDS URL: <http://www.loc.gov/marc/marcdocz.html> (дата обращения: 13.12.2015)

мира в реестре позволяет хранить и обеспечивать массовый доступ пользователей сети к объектам.

Для ориентации пользователя в хранилище, а также для создания программы, осуществляющей поиск по реестру, разработаны идентификаторы.

Идентификаторы – метки, набор метаданных, позволяющий находить искомый объект не методом рекурсивного перебора, а непосредственно с помощью адресных ссылок, ориентируясь на заданные координаты в системе данных. Реестр (хранилище) и идентификаторы сосуществуют в прогрессивной гармонии, направленной на обеспечение доступа к информации и ее охране от посягательств на авторские права. В зависимости от типа хранилища создается идентификатор. Идентификатор может присваиваться одному или нескольким типам объектов, а так же использоваться как в информационном пространстве компьютерной сети, так и на бумажных носителях.

Концептуальный каркас DOI: *indecs Content Model* и *Handle System*

Идеология DOI основана на концептуальных каркасах двух систем, отображающих «социальную инфраструктуру»: Модели Содержательной Индексации МСИ (*indecs Content Model*⁷⁸) и так называемой *Handle System*.

Handle System была создана как часть архитектуры DOA в CNRI, которую основал и возглавляет Роберт Кан⁷⁹, один из создателей протокола TCP/IP при поддержке агентства DARPA и Корпорации Национальная Исследовательская Инициатива (CNRI). Данная система является технологией для назначения, управления и детализации устойчивых идентификаторов (*persistent identifiers*) для объектов любого рода. Специфицированные протоколы позволяют распределенной компьютерной системе хранить идентификаторы (имена или «хэндлы») объектов и раскрывать (представлять) эти иден-

78 Digital Object Identifier System URL: http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (дата обращения: 13.12.2015).

79 Corporation for National Research Initiatives URL: <http://www.cnri.reston.va.us/k-w.html> (дата обращения: 13.12.2015).

тификаторы в информацию о местонахождении, особенностях доступа и прочих качествах использования ресурсов. Эта информация может быть изменена в случае необходимости для отображения текущего состояния дел и/или местонахождения идентифицируемого ресурса без изменения имени («хэндла»).

Настоящая система может структурировать и идентифицировать объекты как сущности первого (базового объектного) класса, в отличие от пакета битов, зависящих от прочих атрибутов, таких как их местонахождение. Иными словами, система позволяет оперировать объектами, имея в основе те же логические идеи, что лежат и в основе объектно-ориентированного программирования. Система была создана как часть более широкого Концептуального Каркаса Сервиса Распределенных Цифровых Объектов⁸⁰, однако используется и в независимых от него приложениях. Содержит опциональные свойства, такие как инфраструктура открытого кода. Система масштабируема для огромного количества номеров объектов без угрозы деградации представляемых данных.

Принципы Handle System:

- Не основана на принципиально изменяемых атрибутах объектов (месте, владельце, прочих атрибутах, которые не могут измениться без изменения референциального тождества)
- Непрозрачность (преимущественно «темные номера»: хорошо известные паттерны могут быть неверно поняты, бессмысленная семантика не будет переводиться на другие языки и вызывать конфликты торговых марок)
- Уникальность в пределах системы (для избегания коллизий и референциальной неточности)
- Обладает опциональными, но удобными поддерживаемыми свойствами (читаемость на естественном языке, совместима с общими системами, например, спецификацией URI)
- Масштабируемость (максимальная нагрузка распределяется на большее количество компьютеров)

80 Wikipedia URL: https://en.wikipedia.org/wiki/Handle_System (дата обращения: 13.12.2015).

- Гибкость (адаптируема к изменениям компьютерной среды, применима в новых приложениях)
- Высокая доверительность (детализация и администрация основана на технических доверительных методах; провайдер имеет полномочия на исключительно длительный период времени)
- Открытая архитектура (подразумевает вовлечение сообщества в разработку инфраструктурных приложений)

Модель Содержательной Индексации была разработана в соответствии с Инициативой ЕС Info-2000, финансируется ЕС и некоторыми организациями, представлявшими издателей, авторов, библиотеки и прочих субъектов в период с 1998 по 2000 годы, использовавших метаданные в своей деятельности, в том числе для электронной коммерции. В словаре системы⁸¹ утверждается, что метаданные для электронной коммерции требуют мультимедийности, мультифункциональности, многоуровневости, многоязычности, транснациональности и мультиплатформенности.

В настоящий момент помимо DOI МСИ используется в следующих системах:

- RDA/ONIX Framework for Resource Categorization
- ISO/IEC 21000-6 (MPEG) Rights Data Dictionary (RDD) – Реестр прав на видеоконтент
- DDEX (Digital Data Exchange) – Система словаря данных и пересылка данных в музыкальной индустрии
- ONIX (Online Information Exchange) – Стандарты для издателей, распространяющих цифровые метаданные о своих продуктах
- Digital Object Identifier System metadata schemes

МСИ обеспечивает взаимодействие по крайней мере в пяти сферах:

- 1) между субъектами медиа-среды (книги, периодические издания, аудиоконтент, аудиовизуальный контент, программное обеспечение, абстракты, визуальный контент);
- 2) между функциями обработки данных (каталогизация, автоматизация документооборота, поиск и управление правами);

81 Digital Object Identifier System URL: http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (дата обращения: 13.12.2015).

- 3) между метаданными от простых до сложных;
- 4) между разным семантическим содержанием разных языков и в рамках одного языка;
- 5) между лингвистическими барьерами разных языков;
- 6) между различными технологическим платформами.

Концептуальный каркас МСИ основан на следующих принципах идентификации:

- *Принцип Уникальной Идентификации* – каждый объект должен быть идентифицирован уникальным образом с помощью имени, фиксирующем его место в пространстве (реальном или виртуальном)
- *Принцип Функциональной Детализации* – должно быть возможно идентифицировать объект в любой необходимый момент.
- *Принцип Установления Авторства* – автор каждого раздела метаданных должен быть надежно определен.
- *Принцип Приемлемого Доступа* – каждый имеет доступ к метаданным, от которых он зависит; конфиденциальность и права на метаданные защищены от третьих лиц, которые от них не зависят.

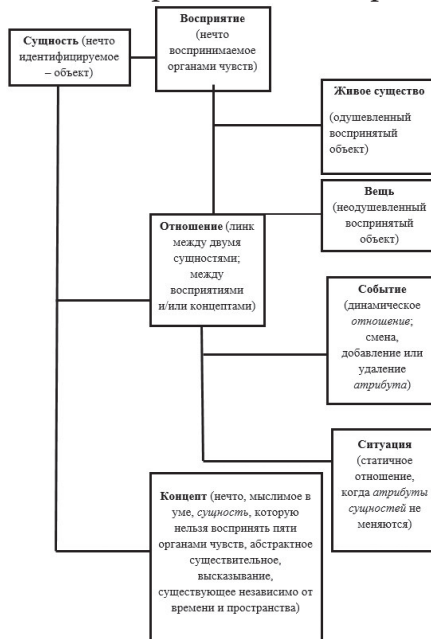


Рис. 13 Онтологическая схема Модели Содержательной Индексации

В данной системе понятий понятие *атрибут* определяется согласно ISO 11179, а понятие *уровня* согласно ISO 11179-3. Словарь содержит также более сотни прочих понятий, таких как *агент*, *лэйбл*, *контекст*, *количество*, *качество*, *материал*, *субъект изменения*, *копия*, *модификация*, *организация* и *многие другие*.

Пример изменения структуры метаданных:

Два человека (А и В) сотрудничали в написании и иллюстрировании книги (X) в Англии в 1999 году. Это может быть выражено как отдельное событие, которое включает: Автора А (агент) + Иллюстратор В (агент) + Книга X (выход (output)) + Англия (контекст) + 1999 (контекст). Однако, если мы хотим отметить, что книга была написана между мартом и августом в Манчестере, а иллюстрации были завершены в октябре в Ноттингеме, нам требуются два отдельных события, поскольку контекстуальные роли более не прилагаются к ролям обеих агентов:

Автор А (агент) + Книга X (выход) + Манчестер (контекст) + Март-Август 1999 (контекст). Иллюстратор В (агент) + Книга X (выход) + Ноттингем (контекст) + Октябрь 1999 (контекст).

Основные приложения DOI в настоящее время включают:

- устойчивое цитирование академических материалов через систему CrossRef – консорциум 3000 издателей;
- исследовательские базы данных, функционирующие через систему DataCite, консорциум ведущих научных лабораторий, технических и информационных провайдеров, научных дата-центров;
- официальные публикации ЕС, распространяемые через EU publications office;
- перманентные глобальные идентификаторы для коммерческого видео-контента, распространяемые через Entertainment ID Registry, более известного как EIDR.

Система идентификаторов DOI

DOI позволяет идентифицировать большой диапазон произведений, таких как изображения, аудио и видео контент, программное обеспечение как в физическом, так и в электронном виде, спектакли, лицензии и т.п. Номера способны отсылать к объектам на различных уровнях и с различной степенью

детализации: номер DOI способен идентифицировать журнал, отдельный выпуск (номер) журнала, отдельную публикацию в журнале или даже конкретный параграф в статье. Выбор уровня детализации лежит на лице, передающем права на интеллектуальную собственность, однако в системе DOI он должен декларироваться как часть метаданных, связанных с именем в системе, на основании соответствующего словаря данных.

DOI – стандарт обозначения выходных данных представленного в сети объекта. Выходные данные – совокупность уникальных данных, содержащих информацию, позволяющую идентифицировать результат интеллектуальной собственности в сети.

Разработка стандарта DOI обусловлена международными правовыми актами, призванными регулировать циркуляцию академической и культурной информации. Однако DOI является уникальным идентификатором, который может быть присвоен различным оцифрованным объектам. Информация, содержащаяся в идентификаторе DOI электронного документа, содержит указатель его местонахождения (например, URL), его имя (название), прочие идентификаторы объекта (например, ISBN для электронного образа книги) и ассоциированный с объектом набор описывающих его данных (метаданных) в структурированном и расширяемом виде. Универсальность или предметная ориентированность идентификатора не означает их автономного существования независимо друг от друга. Напротив, взаимообусловленность идентификаторов между собой позволяет пользователю осуществлять многоуровневый поиск интересующего объекта, а также упростить пользователям доступ к оцифрованным объектам, если они не могут правильно сориентироваться в многообразии представленного контента.

DOI имеет некоторые общие черты со стандартом PURL: наличие указателя местонахождения объекта и его имени (названия).

DOI принят в англоязычной научной среде для обмена данными между учеными. По сути, DOI – это путь к документу в

общем информационно-виртуальном пространстве (как правило, в интернете), для получения необходимой информации⁸².

Идентификатор цифрового объекта представляет собой уникальную строку букв и цифр, состоящую из двух частей: префикс и суффикс. Например, 10.1000/287, где 10.1000 – префикс, или *идентификатор издателя*, составленный из признака идентификатора (10) и строки, указывающей на издателя (1000); 287 – суффикс, указывающий на конкретный объект. Префиксы издателей распределяются регистрационным агентством (DOI Registration Agency). Суффикс формируется издателем и должен быть уникальным у данного издателя. Идентификатор цифрового объекта может объединить существующие идентификаторы, такие как ISBN или SICL. Существенный плюс идентификаторов по типу DOI заключается в возможности предоставления идентификации разными лицами. Пользователь обладает свободой выбора – когда, как, на каких условиях он может получить идентификатор DOI.

Создание идентификаторов и репозиториев цифровых объектов отвечает концепции, заложенной в международных правовых актах, регулирующих правовой режим результатов интеллектуальной деятельности. Всемирная декларация по интеллектуальной собственности от 26.06.2000 гласит: *«В полной мере используя информационные технологии, следует приложить усилия для распространения и расширения знаний общества об интеллектуальной собственности и правах интеллектуальной собственности, с тем, чтобы повысить заинтересованность в интеллектуальном творчестве»*⁸³.

Однако сама суть DOI предполагает идентификацию уже размещенного в сети результата интеллектуального труда. По своей природе оцифрованный объект представляет собой копию воплощенного в реальности результата умственного труда. Таким образом, DOI осуществляет работу

82 URL: <http://crossref.org/> (дата обращения: 13.12.2015).

83 Федеральная служба по интеллектуальной собственности (Роспатент) URL: http://www.rupto.ru/docs/interdocs/worldwide_declaration (дата обращения: 13.12.2015).

исключительно отображением в сети объектов реального мира. Работа с оцифрованным вариантом имеет свою специфику. Субъект способен получить доступ к метаданным оцифрованного варианта объекта, при помощи которых он может найти в реальном мире сам объект. Подобного рода стандарты идентификации позволяют не только свободно ориентироваться в многообразии интеллектуальных объектов, но и предотвращать споры, связанные с установлением авторства относительно каждого объекта. Идентификация оцифрованной версии предполагает указание на идентификационные данные результата умственного труда вне информационно-технологического пространства.

Существует ассоциация компаний, которые осуществляют регистрацию по стандарту DOI, как частному проявлению технологии DOA, в некоторых сферах деятельности. Потребитель самостоятельно знакомится с перечнем услуг, который предоставляет каждая компания (член ассоциации)⁸⁴ в соответствующей сфере деятельности. Регистрация носит добровольный характер и направлена на решение определённых задач, основными из которых являются автоматизация информационных процессов контроля оборота объектов в определенной сфере деятельности, а также защита имущественных и личных прав немущественных прав правообладателей.

Идентификационная сфера DOI ограничена информационным пространством (сеть Интернет). Однако, представляя собой попытку создания максимально полного стандарта идентификации DOI, она включает в себя идентификационные данные непосредственного объекта интеллектуальных прав.

Необходимо обратить внимание на следующие различия между идентификаторами типа DOI и типа ISBN:

1) идентификатор ISBN присваивается только книжным изданиям (узкоспециализированный), в отличие от идентификатора DOA/DOI, который можно использовать для идентификации объекта любой природы;

84 Digital Object Identifier System URL: <http://www.doi.org/faq.html> (дата обращения: 13.12.2015).

2) ISBN предназначена для идентификации тиража издания на бумажном носителе⁸⁵. ISBN де факто применяется для идентификации издания в цифровой форме, но это скорее «натяжка», чем естественное решение;

3) DOI является более универсальным идентификатором, так как представляет собой совокупность любых известных регистрирующей компании идентификационных данных объекта интеллектуальных прав;

4) использование DOI отвечает современным тенденциям, направленным на универсализацию и систематизацию данных (данные о цифровом объекте);

5) имеется еще одна важная особенность, касающаяся цифровых объектов, относящихся к произведениям литературы и, вообще, тому, для выражения чего именно используется язык. Казалось бы, цифровой объект суть электронная версия книги. Но это не так. Сама книга – символическое представление того, что можно произнести на каком-то языке как звуковой системе выражения мыслей. То есть сама книга – цифровой объект, который выполнен с использованием аналоговых технологий (то есть по сути – испачканная чернилами бумага). В отличие от ISBN, технология DOI, с одной стороны, не подразумевает ничего аналогового, с другой стороны, являет собой все ту же фиксацию мыслей, выраженных на каком-то языке, с использованием письменной речи, зафиксированной в электронной форме. Например, при публикации отчета, где для его версии на конкретном языке (английском в данном случае, но есть и на немецком и на французском) указаны 3 разных ISBN – 9789264234024 (PDF); 9789264238718 (EPUB); 9789264232303 (print) (для разных форм представления – в формате PDF, EPUB, а также для печатного экземпляра), а также один код в DOI 10.1787/9789264234024-en⁸⁶.

85 Федеральная служба по интеллектуальной собственности (Роспатент) URL: http://www.rupto.ru/docs/interdocs/worldwide_declaration (дата обращения: 13.12.2015)

86 OECD iLibrary URL: http://www.oecd-ilibrary.org/social-issues-migration-health/indicators-of-immigrant-integration-2015-settling-in_9789264234024-en (дата обращения: 13.12.2015)

Репозитории и реестры (соотношение и виды реестров)

Следует иметь в виду, что идентификация реального объекта (изданной книги, написанной картины и т.д.) и его оцифрованной копии, выложенной в сеть – суть разные сущности (поэтому DOI и ISBN являются *разными онтологиями*). Работа с цифровой версией издания предполагает автоматическое указание на местоположение объекта в реальном мире. *Цифровая версия – объект, отображающий в информационно-технологическом пространстве реально существующий материальный объект.* Возникновение цифровой версии напрямую зависит от существования потенциально копируемого объекта в материальном мире. В то же время, исчезновение из материального мира скопированного объекта может никак не сказаться на онтологическом статусе цифровой версии, находящейся в сети. В зависимости от целей, которые преследует потребитель, исходный объект материального мира может ему совершенно не пригодиться, а копия полностью удовлетворит его пользовательский интерес. Существование различных идентификаторов для цифровых объектов, выложенных в сеть Интернет, является неотъемлемым условием развития информационного пространства и прогресса. Смысл представления объектов в цифровом виде в сети заключается, в том числе, и в возможности одновременно доступа к объекту тысяч и даже миллионов пользователей.

Одним из возможных вариантов повышения качества защиты данных от посягательств является хранение контента в разных хранилищах, возможно, с помощью облачных технологий – необходимое условие стабильности содержательного компонента информационного пространства. Подобное распределение хранилищ повышает шанс на неприкосновенность результатов интеллектуальной собственности. Взлом одного репозитория не позволяет сменить метаданные в реестре и сами объекты для размещения в иных хранилищах. Многие репозитории имеют защиту от подмены. Этот способ реализован через сверку цифровых подписей репозитория и клиента. В случае, когда репозиторий имеет цифровую подпись, а пользовательский компьютер содержит открытый ключ для этого репозитория – такой

репозиторий считается доверенным. При наличии на пользовательском компьютере нескольких подключенных репозиториях, предпочтение отдается доверенным.

Таким образом, репозитории (хранилища) обеспечивают хранение цифровых объектов в сети, они должны быть связаны с реестром (одним или несколькими). Существуют репозитории для хранения программ, написанных на одном языке (например, CPAN для Perl) или предназначенных для одной платформы⁸⁷. Репозитории можно классифицировать:

1. Предназначенный для одного типа объектов.
2. Включающий в себя объекты разных типов.

В России большинство реестров кроме DOI находятся в ведении Роспатента⁸⁸. Российские реестры носят тематический характер, но в то же время составляют единое подмножество в информационном пространстве. Контроль Роспатентом основных используемых реестров в стране обоснован тем, что в сфере государственной компетенции находится регулирование патентных правоотношений. Согласно ст. 1353 ГК РФ, «исключительное право на изобретение, полезную модель или промышленный образец признается и охраняется при условии государственной регистрации соответствующих изобретения, полезной модели или промышленного образца, на основании которой федеральный орган исполнительной власти по интеллектуальной собственности выдает патент на изобретение, полезную модель или промышленный образец»⁸⁹. Каждый реестр, представленный Роспатентом, отвечает определенной тематике, что позволяет осуществлять быстрый поиск необходимых объектов. В государственном ведении находятся следующие реестры:

1. АБС ФГУ ФИПС (Автоматизированная библиотечная информационная система Федерального государственного

⁸⁷ URL: <https://help.ubuntu.com/community/Repositories> (дата обращения: 13.12.2015)

⁸⁸ Федеральная служба по интеллектуальной собственности (Роспатент) URL: <http://www.rupto.ru/about/fgis> (дата обращения: 13.12.2015)

⁸⁹ КонсультантПлюс URL: https://www.consultant.ru/popular/gkrf4/79_10.html (дата обращения: 13.12.2015)

учреждения «Федеральный институт промышленной собственности» Федеральной службы по интеллектуальной собственности, патентам и товарным знакам.)

2. АС ЭГР (Электронные Государственный реестр товарных знаков и знаков обслуживания Российской Федерации, Государственный реестр наименований мест происхождения товаров и Перечень общеизвестных в Российской Федерации товарных знаков.)⁹⁰

3. ФГИС АИС учета РИД (Федеральная государственная информационная система учета результатов научно-исследовательских, опытно-конструкторских и технологических работ военного, специального и двойного назначения, права на которые принадлежат Российской Федерации.)⁹¹

Поиск и хранение информации осуществляется за счет соединения реестра и репозитория в единую информационную систему, направленную на каталогизацию, систематизацию, а также хранение и защиту информации о представленных в сети цифровых объектов.

DOI как «система вещей» – претендующая на полноту онтология объектов и предметной деятельности

Концептуальный каркас, лежащий в основании DOI, основан на Модели Содержательной Индексации и Handle System. Данные системы являются мощными онтологиями, описывающими динамические и статические отношения между авторами, правообладателями, материальными, цифровыми объектами, данными и метаданными. Можно утверждать, что DOI является, выражаясь языком французского философа М. Фуко, полноценной «системой вещей», имеющих или потенциально способных иметь значение для человека в его предметной деятельности.

⁹⁰ Федеральная служба по интеллектуальной собственности (Роспатент) URL: <http://www.rupto.ru/about/fgis> (дата обращения: 13.12.2015)

⁹¹ РосКомНадзор URL: <http://rkn.gov.ru/it/register/#> (дата обращения: 13.12.2015)

Система ISBN также является онтологией, однако гораздо более слабой, предмет которой ограничен книжной продукцией. Эта система является не полноценной «системой вещей», а «системой книг». Следует иметь в виду, что система (сеть) книг сама по себе является информационным пространством, описывающим систему вещей – *deskриптивным и концептуальным уровнем по отношению к объектному уровню материальной реальности*, референтом которой она является в печатных текстах. До появления цифровых носителей информации она отвечала требованиям хранения и репрезентации знания человечества. Однако, к настоящему моменту времени эта функция отошла на второй план.

Более того, *появление информационной надстройки в виде Глобальной Сети переводит систему печатной продукции из метауровня представления знания на объектный уровень по отношению к своим репрезентационным возможностям.*

Следует иметь в виду, что система DOI в достаточно широких рамках *нечувствительна к различиям между системами категоризации материальных объектов*. Приведенная система категорий основана на различии видов человеческой деятельности, которая носит эмпирический характер (в т.ч. исторически сложившийся) и не вправе называться полноценной классификацией с точки зрения требований формальной логики. Она, скорее, отображает устоявшиеся представления о категориях материально-предметной и духовной деятельности человека. В принципе, система категорий может быть и иной. Она является вторичной по отношению к системе категорий, заложенной в онтологии МСИ и Handle System, поскольку именно они являются несравненно более мощными с логической – формальной и содержательной – точек зрения. Это означает, что рецепция DOI постепенно приведет к коррекции категоризации объектов материального мира в сторону большей логической и динамической упорядоченности в прагматических целях.

В основе онтологии DOA/DOI лежит фундаментальное теоретико-познавательное различие эмпирическими данными пяти органов чувств и сверхчувственным концептуальным знанием. В настоящий момент пять чувств могут быть редуцированы к

четырем, поскольку запахи не используются для представления цифровых объектов. Именно эпистемический способ доступа к экземплярам, образам или репрезентациям цифровых объектов является тем существенным признаком, который должен лежать в основе систематизации референтов образов, т.е. материальных объектов в цифровом пространстве.

Преимущества системы DOI

Принципиальные преимущества:

- *Твердая референция по отношению к агентам предметной деятельности.* Основана на *внутренних атрибутах* объектов (месте, владельце, прочих атрибутах, которые не могут измениться без изменения референциального тождества)
- *Защищенность от семантической нагруженности естественным языком.*
- *Непрерывная динамическая идентификация, устойчивая к смене внешних атрибутов объектов.*
- *Практически неограниченный потенциал гипертекстуальности* – повышения связности цифровых объектов друг с другом.
- *Применимость для локализации как объектов материального мира в физическом пространстве, так и для локализации цифровых объектов в пространстве глобальной сети.*
- *Независимость как от онтологического типа объекта (материальный, цифровой), так и его содержательного типа (текст, аудио, видео и т.п.).*

Технические преимущества:

- *Практически неограниченный потенциал нумерации.*
- *Уникальность идентификаторов в пределах системы.*
- *Способность отображаться на естественных языках.*
- *Избыточность хранения метаданных.*
- *Масштабируемость.*
- *Гибкость.*
- *Открытая архитектура.*

Коммуникативный потенциал DOI

Данная система идентификаторов, реализующая принципы мультимедийности, мультифункциональности, многоуровне

ности, многоязычности, транснациональности и мультиплатформенности, позволяет увеличить уровень связности данных, преодолевая барьеры:

- между типами медиа (книги, периодические издания, аудио контент, аудиовизуальный контент, программное обеспечение, абстракты, визуальный контент);
- между функциями обработки данных (каталогизация, автоматизация документооборота, поиск и управление правами);
- между уровнями метаданных;
- между семантическими и лингвистическими структурами (языками на разных уровнях организации);
- между различными технологическим платформами.

Инфраструктурные преимущества

1) В самой основе системы DOI лежит идеология «миграции» знания в глобальную сеть, она адаптирована именно для этих целей. Тренд на переход на электронный документооборот, развитие и внедрение идеологии Электронного Правительства повлечет за собой внедрение, по крайней мере, системы подобного типа, сопоставимой по своим онтологическим возможностям.

2) Тренд на миграцию данных в облачные хранилища демпфирует проблему корыстной порчи и подмены метаданных, в т.ч. проблему соблюдения авторских прав, т.к. он облегчит контроль на этапе загрузки с лицензированных операционных систем, предустановленных на мобильных устройствах. В случае усиления этого тренда возможна ситуация, когда модификация метаданных объекта станет очень сложной задачей вследствие того, что метаданные станут сложными по своей структуре и будут спрятаны среди прочих данных. Рост сложности программных продуктов, приближение их к возможностям искусственного интеллекта влечет за собой естественные формы поведения цифровых объектов, рост связности метаданных, в результате чего их свойства начинают приближаться по сложности к свойствам естественных объектов, трудно поддающихся подделке.

3) Возможна ситуация, когда произойдет отход от привычной системы хранения данных, основанных на файловых системах,

по крайней мере, на уровне репрезентации цифровых объектов конкретным пользователям. Это возможно благодаря миграции данных в облачные хранилища, поскольку значение метаданных будет расти, а объектных данных снижаться по отношению к операциям обработки информации. Даже если конкретный цифровой объект перестанет быть конкретным файлом, он станет не более чем набором данных, хранящихся «неизвестно где», а также метаданных, представляющих собой атрибуты объекта и соответствующие ссылки для доступа к нему. Даже в случае изменения способа хранения объектных данных, система метаданных имеет существенные шансы сохраниться, т.к. она является онтологически более мощным концептуальным каркасом – более высокого уровня абстракции, нежели система форматов хранения объектных данных.

Необходимость национального российского регистратора

В Российской Федерации услугами по предоставлению идентификатора DOI занимается Некоммерческое партнерство «Национальный Электронно-Информационный Консорциум» (НЭИКОН) (с 20 августа 2014).

НЭИКОН формально определяется как консорциум российских библиотек, занимающийся организацией подписки на электронную научную информацию на территории России⁹².

С 20 августа 2014 года НЭИКОН получил статус первого в России официального партнера и агента Международной ассоциации по связям издателей (Publishers International Linking Association, Inc. – PILA), корпорации, являющейся владельцем и оператором базы данных цифровых идентификаторов CrossRef. Кроме того, НЭИКОН является официальным агентом (Sponsoring Entity) Crossref в России и имеет полномочия осуществлять оплату членства организаций-соискателей в «Международной ассоциации по связям издателей» (Publishers International Linking Association, Inc. – PILA) и оплату цифровых идентификаторов DOI. Префикс DOI и идентификаторы принадлежат издательству (или изданию). Издательство

⁹² НЭИКОН URL: <http://neicon.ru/ru/advances> (дата обращения: 13.12.2015).

самостоятельно формирует идентификаторы и загружает их и метаданные статей в систему Crossref.

Следует обратить внимание, что PILA рассчитывает сумму ежегодного взноса по календарному году (1 января-31 декабря) и взимает оплату по двум направлениям:

1) Ежегодный членский взнос в PILA (уплачивается единовременно, один раз в год, например, в январе; взнос привязан к курсу доллара США и, в обычном случае, равняется \$ 300 в год для 1-4 журналов и \$ 600 в год для 5 и более журналов (в пересчете на рубли на дату заключения договора). Размер регистрационного сбора сокращается пропорционально количеству оставшихся до конца года кварталов.

2) Оплата каждого DOI, которая осуществляется по факту загрузки DOI на основе ежеквартального счета. При работе через НЭИКОН стоимость одного doi составляет 120 руб. для новых выпусков и 75 руб. для архивных выпусков (по состоянию на 2015 г. архивными считаются выпуски старше 2013 г.). Т.е., например, в I квартале года вы загрузили 100 новых DOI и, соответственно, в начале 2 квартала вы получите от НЭИКОН счет на 12.000 руб., акт, счет-фактуру и т.д.

НЭИКОН не является независимой организацией, выполняя агентские функции иностранного партнера. Существенным в данной ситуации является коммерческий характер регистраторов, поскольку это влечет за собой зависимость от тарифной политики, которая неизбежно транслируется на потребителей. В условиях нестабильности экономической обстановки, высокой инфляции колебания тарифов могут существенно варьироваться, негативно сказываясь на таком важном параметре, как доля действительно присваиваемых идентификаторов из всех возможных.

Тем не менее, кардинальный способ разрешения существующей ситуации в целях создания национальной системы идентификации произведений в сфере искусства, науки и образования заключается не в увеличении количества регистраторов и не в государственных субсидиях таковым, а в создании собственного регистратора высокого уровня и,

возможно, *специальных идентификаторов* в рамках Модели Содержательной Индексации (на которой основана DOI) для специальных нужд.

Данная задача может быть решена либо созданием специального ведомства на правах федерального агентства или федерального унитарного предприятия либо наделение полномочиями регистратора существующих организаций, обладающих сетью представительств в субъектах федерации. В качестве такой сети может выступать, например, сеть библиотек либо иной государственной организации, обладающих региональной сетью представительств.

Общие принципы:

1. *Приоритетность национального регистратора для нужд реализации государственной политики в области культуры, науки и образования.* Организации, обладающей правами национального регистратора, вменяется бесплатное или условно-бесплатное (за счет бюджетного финансирования и/или с отложенной оплатой) присвоение DOI-идентификаторов и прочих идентификаторов для таковых нужд. Такое введение должно осуществляться с целью устранения препятствий для опубликования в сети цифровых объектов их авторами. В противном случае, авторы, не обладающие, временно или вообще, возможностью оплатить присвоение DOI и аналогичных ему идентификаторов, рискуют выпасть из зоны максимальной пользовательской доступности. Эпоха массовой оцифровки данных и их систематизации не предусматривает такого отношения к авторам, чьи материальные возможности не позволяют оплатить идентификационный номер.

2. *Наделение национального регистратора функциями ведения реестра.*

3. *Сохранение возможности функционирования существующих и создания альтернативных регистраторов для коммерческих нужд на конкурентной основе.*

4. *Диверсифицируемость идентификаторов.* Данный принцип подразумевает создание специальных национальных идентификаторов по аналогии с США, где существуют собственные

идентификаторы одновременно с DOI (идентификатор Национальной библиотеки США). Массив информации, зафиксированной при помощи идентификатора DOI, огромен. В случае, если пользователь осуществляет поиск статьи на территории Российской Федерации, и он точно уверен в том, что издание было опубликовано в России, то поиск по национальному идентификатору может сократить время поиска. Национальный идентификатор может в более короткий срок помочь найти идентифицируемый объект в случае, если произойдет сбой в идентификационной системе DOI. Национальные идентификаторы могут позволить создать базы данных с закрытым или ограниченным доступом к информационным объектам, доступ к которым может или должен быть ограничен в связи с их важностью для интересов государственной безопасности.

5. *Диверсификация репозиториев* в целях облегчения доступа к данным, безопасности данных от повреждений и киберугроз. В качестве операторов репозиториев могут выступать телекоммуникационные компании, обладающие собственной высокоскоростной телекоммуникационной сетью и сетью дата-центров в различных регионах России. Возможна реализация хранения данных в облачном виде на основании национальной облачной платформы.

6. *Возможность распространения услуг по предоставлению идентификаторов на дружественные государства СНГ.*

Национальный регистратор DOI

Запуск системы регистрации DOI должен осуществляться некоммерческими агентами, которые нацелены на развитие науки, образования и культуры в России и зарекомендовали себя в этой сфере. Предполагается, что партнерами запуска DOI должны стать: Фонд поддержки интернет, Ассоциация интернет-издателей, Ассоциация «Открытая наука», Консорциум НЭИКОН, Российская государственная библиотека, Государственная публичная научно-техническая библиотека РФ.

На стартовом этапе предполагается выделение номерной емкости в размере 50 миллионов идентификаторов с наделением уполномоченных агентов (через регистратора, например, Фонд

поддержки интернет) правом выдавать идентификатор первоисточникам и другим материалам из оцифрованных коллекций, предназначенных для публикации бесплатно (при условии публикации полного текста в открытом доступе).

В частности, предполагается, что номера DOI будут получать все документы, опубликованные в рамках НЭБ, библиотеки диссертаций и других публичных коллекций РГБ, все материалы, опубликованные в открытом научном архиве ГПНТБ РФ, все дипломные и учебные работы и исследования, опубликованные в открытых архивах и депозитариях (таких как «Википедия», «КиберЛенинка», «Научный Корреспондент» и «Вернский», а также консорциума НЭИКОН – для материалов, предназначенных к публикации в открытом доступе на территории РФ).

В Российской Федерации DOI должен выдаваться бесплатно при условии публикации полного текста произведения в открытом доступе. Объем условно-бесплатной номерной емкости идентификаторов должен быть достаточным для публикации всех оцифрованных коллекций РГБ и ГПНТБ РФ, а также материалов открытых архивов и депозитариев. Выделение указанной номерной емкости запланировано на следующем этапе развития проекта.

Массовое принятие DOI в качестве стандарта идентификации для произведений науки и культуры, в т.ч. первоисточников, создаст условия для разворачивания системы платной регистрации, в т.ч. для тех авторов, которые не готовы публиковать полные тексты своих документов.

Наличие двух так важных инфраструктурных проектов, как Национальная электронная библиотека РГБ и научный архив и депозитарий ГПНТБ РФ позволяет предложить следующий вариант формирования электронной инфраструктуры выдачи DOI.

Уже зарекомендовавшие себя публичные репозитории с открытым доступом, такие как «Библиотека Мошкова», «Википедия», «Научный Корреспондент», «Вернский», «КиберЛенинка» смогут получить возможность представить в научный архив и НЭБ произведения, которые с их точки зрения и с точки зрения их пользователей: 1) являются удостоверенного качества;

2) действительно принадлежат авторам, которые указаны в качестве таковых. Кроме того, ресурс должен обладать правом на размещение указанных произведения. В случае, если произведения отвечают перечисленным требованиям, они могут получить DOI при размещении публикации в открытом доступе в научном архиве или в НЭБ. При этом, скорее всего, с этой целью может быть сформирована дополнительная подсистема, оператором которой может быть консорциум НЭИКОН, АИИ и Фонд поддержки интернет. Предполагается, что подготовка к размещению в НЭБ и научном архиве будет осуществляться через предварительный этап публикации произведений, в процессе которого авторы могут предъявить претензии (например, если у опубликованных сиротских произведений будут найдены правообладатели, а также в случае, если в репозитории размещена недостоверная или некачественная копия). С этой целью в рамках проекта «Общественное достояние» предусмотрено создание реестра произведений в режиме общественного достояния, который можно дополнить реестром произведений, которые распространяются в режиме открытого доступа, т.е. на условиях открытых лицензий.

Предварительным идентификатором на предварительном этапе станет UUID, реестр которых сформирует каждый архив и репозиторий. UUID будет присваиваться с префиксом, в котором в определенной форме должна быть обозначена площадка, а затем полученный идентификатор сформирует, по сути, тот DOI, который будет выдан при размещении в НЭБ и научном архиве. В итоге DOI будет состоять из следующих элементов: национального префикса (собственной DOI) и суффикса, в котором указаны: 1) площадка-источник – в качестве префикса к UUID, который объясняет происхождение файла и документирует его, и собственно UUID произведения.

Хотелось бы особо обозначить важность качественного подбора партнеров для запуска системы DOI в России. По сути, перечисленные выше проекты принадлежат частным лицам или организациям, представленным в коротком списке в начале этого раздела. Поскольку важна именно подтвержденная временем

репутация сторонников открытого доступа к знаниям, список доверенных репозиториев может быть расширен и далее за счет включения в него изданий и проектов, размещающих материалы в открытом доступе на условиях открытых лицензий.

Сформированная таким образом система авторизованных агентов внедрения DOI станет частью новой инфраструктуры системы знаний – своего рода федеральной резервной системы. Однако в отличие от США в нашем случае речь идет не о деньгах, а об открытом доступе к документам и знаниям. А банки – это не кредитные организации, а банки знания – «вход» в национальную систему в виде НЭБ и базы научных первоисточников: курсовых и дипломных работ, статей, диссертаций, тезисов, презентаций и рецензий.

Раздел 4.

Исследование использования системы DOI (Digital Object Identificators) и построение пилотного проекта для работы с данной системой

Технологии Digital Object Architecture (DOA) и Digital Object Identification (DOI, правильный перевод – «цифровая идентификация объектов») как ее частный случай широко применяются для идентификации объектов в различных системах. Впервые представленная в работе 1995 года «A Framework for Distributed Digital Object Services» (<http://www.cnri.reston.va.us/k-w.html>) Роберта Кана и Роберта Виленски (Robert Kahn, Robert Wilensky), она получила широкое распространение и использование для идентификации и ссылок на различные объекты, как цифровые, так и материальные.

Несмотря на востребованность в международном сотрудничестве и удобство для использования, технология на данный момент не получила должного распространения в России. В данном разделе рассматриваются результаты исследования как самой технологии, так и перспектив внедрения ее для решения разного рода задач идентификации объектов. Исследование проведено российскими специалистами и учитывает особенности внедрения DOI в отечественной практике.

Проведение исследования складывалось из двух этапов – ознакомление с имеющимся материалом и установка, настройка и апробация ПО.

На первом этапе было проведено ознакомление с публично доступной информацией (сайты <http://www.doi.org>, <http://www>.

handle.net, <http://cordra.org>, <http://eidr.org> и др.). По результатам этого этапа была осознана схема взаимодействия компонентов системы, обеспечивающей применение DOI в информационных системах.

На втором этапе, после получения Фондом содействия развитию интернета «Фонд поддержки интернет», который являлся заказчиком данного исследования, префикса 10.10329 в системе Handle, была произведена установка и настройка ПО.

Общее описание акторов инфраструктуры DOA/DOI

Корневой организацией для управления системой DOA/DOI является DONA Foundation (<https://www.dona.net/>). Она управляет пространством имен Handle и поддерживает GHR на базе Multi-Primary Administrators (MPA).

Надо отметить, что система отношений между различными субъектами в рамках инфраструктуры Handle/DOI окончательно не сформирована (так, на данный момент International DOI Foundation не является MPA в рамках DONA, хотя и ведет по этому поводу активные переговоры). Обратной стороной этого является отсутствие документированных схем глобального взаимодействия иначе как с непосредственными контрагентами.

Корневой структурой в рамках DOI (Handle prefix 10) является International DOI Foundation (IDF), сайт <http://www.doi.org/>. Предусмотрено членство организаций в IDF. Из 4 вариантов членства наиболее интересны 2: обычное членство (General Membership) и членство в виде Registration Agency (RA). Обычное членство предусматривает работу в системе DOI, членство в качестве RA подразумевает работу по регистрации объектов для конечных пользователей.

В рамках подсистемы регистрации системы DOI иерархическая структура и взаимоотношения контрагентов четко прописаны. В системе DOI регистрируются регистрационные агентства (Registration Agency, RA), каждое из которых действует в

своей предметной и/или территориальной области, определяет свою модель данных и оказывает услуги конечным пользователям, напрямую или через агентов

Требования к RA

В рамках системы DOI к RA предъявлен ряд требований, описывающих политику RA как акторов в рамках системы DOI. Эти требования изложены в DOI Handbook (http://www.doi.org/doi_handbook/8_Registration_Agencies.html).

Базовые технические принципы взаимодействия и полномочий в рамках DOI:

- RA определяет и проводит свою политику, не противоречащую требованиям IDF;
- регистрирует префиксы в GHR в необходимом объеме;
- должно поддерживать качество услуг:
- аккуратность данных;
- консистентность метаданных;
- права доступа и управления;
- использует программное обеспечение, предоставленное CNRI. В CNRI поднимается зеркало LHR RA. Эта деятельность CNRI определяется договором о технологическом партнерстве с IDF.

RA имеет право оказывать дополнительные услуги, не специфицированные в договоре с DOI. RA может оказывать свои услуги, как по принципу юрисдикции, так и по отраслевому признаку. Описание юридических и финансовых требований к RA находится за рамками данного отчета.

Взаимодействие конечных пользователей с системой DOI

Администраторы отдельных объектов регистрируют объекты, обращаясь к RA, сфера деятельности которого соответствует профилю зарегистрированных объектов. При этом об идентифицируемом объекте сообщается информация, специфичная для агентства и категории объекта в соответствии с используемым профилем метаданных.

Администраторы отдельных объектов могут менять информацию о своих объектах (кроме непосредственно префиксов), поддерживая актуальность указания на объекты.

Администраторы объектов могут передавать права на управление объектами другим лицам в соответствии с процедурами передачи внутри RA, специфицированными этим RA. При этом услуги, оказываемые этими RA, могут отличаться.

Администраторы объектов могут передавать управление своими идентифицируемыми объектами от одного RA к другому, если сфера деятельности принимающего RA позволяет осуществлять такие операции.

Конечные пользователи могут использовать публичные сервисы (<http://hdl.handle.net/>, <http://dx.doi.org/>) для резолвинга идентификаторов. При этом предоставляется только та метаянформация, которая доступна для публичного доступа.

Модель данных DOI и ее взаимоотношения с системой Handle

Сама по себе система Handle предоставляет довольно бедный базовый набор метаданных, но предусматривает возможность расширения. Так, есть служебный префикс 10320, предусматривающий регистрацию новых типов (сейчас используется только тип `hdl:10320/loc`, отвечающий за многовариантный резолвинг `hdl`-идентификатора в URL в зависимости от страны и запрашиваемого языка). Тем не менее, базовая система предусматривает возможность установления связей между отдельными DOI-идентифицированными объектами. Это позволяет описывать практически произвольные отношения между объектами, идентификаторы которых есть в системе DOI.

Система метаданных DOI специфицирована стандартом ISO 26324. DOI предписывает более широкий базовый набор метаданных, позволяющий классифицировать и точнее описать объекты по ряду признаков. При регистрации объектов в DOI RA специфицирует еще более широкий профиль, специфичный для своей деятельности. При резолвинге DOI-идентификатора средствами Handle клиент может попасть на сайт RA, где метаданные будут представлены в человекочитаемом виде, или получить их в машиночитаемом виде через API. Более подробное описание подхода DOI к метаданным описано в DOI Handbook (http://www.doi.org/doi_handbook/4_Data_Model.html). В частности, устанавливается

классификация по способам восприятия объекта (зрение, осязание...), данные о связанных сторонах, ссылки из других систем именования (ISBN, ISSN) и др. Подробное описание базовой модели метаданных DOI приведено в других разделах книги.

RA могут устанавливать и использовать свои профили метаданных и давать рекомендации по их использованию. Примеры такого рода рекомендаций можно посмотреть на сайте EIDR (<http://eidr.org>). Для установления соответствия и обеспечения интероперабельности между профилями метаданных различных агентств, а также не входящими в систему DOI профилями метаданных (ONIX, DDEX) используется Vocabulary Mapping Framework (VMF).

В рамках International DOI Foundation предусмотрены процедуры изменения профиля метаданных.

Проведение исследования

Использование ПО от CNRI

В соответствии с условиями членства в DOI Foundation для работ с системой Handle используется серверное ПО от CNRI. Это ПО состоит из трех логических компонентов, объединенных в два программных продукта.

1) Система резолвинга (Handle). ПО доступно для скачивания с сайта <http://www.handle.net>;

2) Реестр цифровых объектов. ПО доступно для скачивания с сайта <http://www.cordra.org>, объединено с п.3;

3) Репозиторий цифровых объектов. ПО доступно для скачивания с сайта <http://www.cordra.org>, объединено с п.2.

ПО написано на языке Java. ПО предоставляет JSON API и web-интерфейсы для выполнения ряда задач.

В соответствии с условиями членства в IDF, каждая организация-член имеет право на выделение Handle-префикса в тестовых целях. В рамках выполнения договора был получен и использован в экспериментах префикс 10.10329.

Была выполнена инструкция по установке ПО Handle, в результате чего был поднят тестовый hdl-сервер, работающий с

выделенным префиксом. Была произведена регистрация сайта Фонда в системе Handle в качестве объекта типа URL. Созданный идентификатор – doi:10.10329/fondpiinfo.

Была выполнена инструкция по инсталляции и настройке ПО Cordra. ПО было установлено на той же машине, что и ПО Handle, что позволило настроить инсталляции для совместного использования.

Настроенное ПО позволило произвести регистрацию ряда объектов в системе DOI и оценить объем работ (аналитические работы и программирование), необходимых для построения системы, пригодной для промышленной эксплуатации.

Рамочное описание функциональности регистратора

Проведенное исследование выявило примерный объем работ по проектированию и программированию системы, предназначенной для реализации функциональности регистратора объектов в системе DOA/DOI, а также управления (администрирования) данными об объектах. Под управлением данными об объектах здесь понимается модификация метаданных объектов имеющими соответствующие полномочия акторами. Ядром системы регистратора предполагается распространяемое CNRI ПО Cordra и Handle Server.

Для построения полнофункциональной системы осуществления деятельности RA потребуется спроектировать следующие дополнительные логические и программные компоненты:

1) Одна или несколько моделей данных, актуальных для предметной области/областей, в которой регистратор планирует регистрировать объекты. Такие модели данных задаются схемами в формате JSON и добавляются в систему реестра/репозитория. Модель данных может предусматривать ссылки на альтернативные системы идентификации объектов (ISBN, ISSN, ISNI, ORCID и пр.) для упрощения управления данными;

2) Административные web-интерфейсы для управления имеющими соответствующие полномочия операторами системы метаданными в реестре объектов и объектами в репозиториях, просмотра истории изменений, журналов действий и прочей служебной информации;

3) Web-интерфейсы для клиентов, имеющих договорные отношения с регистратором. Через эти интерфейсы клиенты получают возможность управлять «своими» идентификаторами, метаданными и идентифицируемыми объектами. Эти интерфейсы должны предусматривать систему прав, характерную для выбранной предметной области. Для доступа в эти web-интерфейсы должна требоваться авторизация;

4) API, представляющая клиентам возможность осуществления операций со «своими» идентификаторами, метаданными и идентифицируемыми объектами в автоматическом режиме. API должна принимать команды в одном из распространенных форматов (JSON, XML) и допускать массовую обработку объектов. Для осуществления операций через API должна требоваться авторизация. Реализация API не является необходимым компонентом для запуска RA, поэтому может быть отложена на второй этап развития;

5) Подсистема проверки регистрируемых объектов на уникальность. Целью этой подсистемы является выявление и информирование о потенциально повторной регистрации одних и тех же объектов. Система должна учитывать метаданные, относящиеся к вновь идентифицируемому или модифицируемому объектам, и, возможно, другие обстоятельства;

6) Подсистема нотификации. Эта подсистема должна обеспечивать клиентам оповещение об изменениях в администрируемых ими объектах;

7) Публичные интерфейсы. Эти интерфейсы предназначены для доступа неавторизованных пользователей и просмотра/поиска данных в реестре/репозитории по выбранным критериям. Для просмотра через публичные интерфейсы должны быть доступны только открытые атрибуты/объекты, зарегистрированные в системе.

Вероятно, при использовании компонентов системы от CNRI придется решать вопрос о локализации (русификации) интерфейсов для облегчения работы с ними.

Система должна предусматривать стандартное техническое обслуживание эксплуатирующей стороной (системное администрирование, резервное копирование, репликация). Система

эксплуатации должна предусматривать мониторинг нормально-го функционирования компонентов (доступность web-интерфейсов, API, базы данных и пр.).

Система должна предусматривать распределенность как самого реестра/репозитория, так и доступа к реестру/репозиторию, что позволит в дальнейшем справляться с ростом нагрузки.

Кроме вышеперечисленного, должна быть подготовлена договорная основа для взаимодействия с контрагентами, специфицирующая права и обязанности сторон по отношению друг к другу. Должны быть разработаны правила пользования системой, а также формальные процедуры разрешения конфликтов, связанных с нарушениями правил пользования системой.

Регистратор должен иметь политику назначения новых префиксов, как верхнего уровня, так и производных, для более четкого позиционирования регистрируемых объектов.

Сама по себе описанная система могла бы предоставить технологическую основу для построения таких сервисов, как реестр объектов интеллектуального права. В частности, получение доступа через систему DOI к объектам позволило бы учитывать реальную востребованность тех или иных объектов, что может быть положено в основу разнообразных форм бизнеса.

Ориентировочный план реализации функциональности RA*

**В таблице ниже представлен приблизительный план реализации функциональности RA, организационные компоненты разработки не перечислены.*

Задача	Этап	Требуемые ресурсы	Примечание
Разработка детализированного ТЗ на систему в целом	0	2 месяца. 2 человека (менеджер проекта, исполнитель)	
Разработка профиля метаданных в соответствии с предметной областью	0	2 месяца на первый профиль, 1,5 месяца на каждый последующий. 2 человека (аналитик, представитель заказчика)	Может выполняться параллельно с предыдущей задачей

Инсталляция и настройка ядра системы (Cordra и Handle.net registry)	1	2 недели, 1 человек	Может выполняться на предварительном этапе
Создание web-интерфейсов для клиентов	1	3 месяца, 2 человека	
Создание административных web-интерфейсов к ядру системы	1	2 месяца, 2 человека	
Создание публичных web-интерфейсов	1	2 месяца, 2 человека	
Создание системы нотификации	1	1 месяц, 1 человек	
Реализация резервного копирования и обеспечение отказоустойчивости системы, в том числе мониторинга	1	2 месяца, 2 человека	
Ручное тестирование	1	1 месяц, 1 человек	Может частично выполняться параллельно с разработкой
Создание API для массовых операций (включая автоматические тесты)	2	2,5 месяца, 1 человек	
Система проверки регистрируемых объектов на уникальность	2	1 месяц, 1 исполнитель и 1 аналитик	Впоследствии система должна подвергаться доработке при внедрении новых профилей метаданных
Тестирование	2	2 недели, 1 человек	

Таким образом, для создания системы необходимо как минимум около 37 человеко/месяцев, при этом трудозатраты для

разработки и запуска системы в эксплуатацию (этапы 0 и 1) оцениваются примерно в 30 человеко/месяцев. Длительность работ по этапам 0 и 1 с учетом возможной параллельности проведения некоторых из них оценивается не менее чем в 7 месяцев.

Стоимость разработки и пуска системы в эксплуатацию составляет по предварительной оценке от 7 до 10 миллионов рублей без учета стоимости оборудования и административно-управленческих расходов (на аренду помещения, бухгалтерию, связь и пр.)

Ориентировочное количество человек для эксплуатации системы без дополнительных доработок – 2 человека с частично дублируемой функциональностью (системный администратор и программист).

Оценка требуемых аппаратных ресурсов

Для инсталляции и эксплуатации системы Handle требуются следующие аппаратные ресурсы:

- CPU: минимум single core 1GHz, рекомендовано single core 2GHz;
- оперативная память: 2 Гб минимум, 4 Гб рекомендовано;
- жесткий диск: 100 Мб необходимы для инсталляции, 20 Гб рекомендованный минимум.

Для инсталляции и эксплуатации Cordra требуются следующие аппаратные ресурсы:

- CPU: минимум single core 1GHz, рекомендовано single core 2GHz;
- оперативная память: 500 Мб минимум, 2 Гб рекомендовано;
- жесткий диск: 200 Мб необходимы для инсталляции, потребность в дополнительном пространстве зависит от того, используется ли только функциональность реестра или функциональность репозитория объектов.

Исходя из соображений резервирования, необходимо 2 машины под Cordra/Handle и 2 машины под web-интерфейсы. Кроме того, требуется дисковое пространство под резервные копии данных. Оптимальным было бы приобретение или инсталляция нескольких виртуальных машин на облачном хостинге с возможностью приобретения дополнительных ресурсов при росте нагрузки на систему.

ЗАКЛЮЧЕНИЕ

Технология DOA/DOI позволяет осуществлять однозначную персистентную идентификацию объектов, в которой заинтересованы правообладатели этих объектов. Это открывает широкие перспективы для системы в России в случае появления субъекта, наделенного полномочиями регистратора в системе DOA/DOI. Широкое присвоение идентификаторов и популяризация их использования может понизить входной порог для внутрироссийских акторов для выхода на международную арену.

Конкурентным преимуществом для российского RA может быть упрощение системы взаимодействия с внутрироссийскими структурами по сравнению с международными структурами (и даже их российскими представительствами) и отказ от необходимости вести валютные расчеты. Кроме того, отношения с российским регистратором будут ближе к повседневной бизнес-практике заинтересованных участников отношений.

Эти обстоятельства вкупе с проникновением DOI-идентификаторов в повседневные практики многих информационных процессов делают целесообразным развитие применения технологии DOA/DOI в России и образование российских регистраторов.

ОГЛАВЛЕНИЕ

Аннотация.....	2
Введение.....	3
Раздел 1. Модели идентификации цифровых объектов в интернете: стандарты и перспективы	5
Цифровые объекты.....	5
Идентификация объектов в интернете.....	6
Метаданные и требования к описанию объектов.....	10
Системы идентификации объектов.....	13
Модели идентификации цифровых объектов на основе децентрализованного распределенного реестра.....	38
Системы DOA/DOI.....	43
Идентификация объектов в контексте интернета вещей.....	66
Раздел 2. Реестр цифровых объектов как инструмент идентификации контента в цифровую эпоху	73
Особенности правового регулирования гражданско-правового оборота объектов авторских прав в сети Интернет.....	73
Возможности реформирования института охраны авторских прав в России.....	90
Реестр цифровых объектов как востребованный инструмент фиксации метаданных о произведении и правообладателе.....	103
Принципы функционирования реестра как системы идентификации контента: унификация, открытость, возмездность.....	115
Анализ мировой практики по созданию и функционированию цифровых реестров (на примере архитектуры DOI).....	117
Раздел 3. Современные инструменты регистрации и идентификации в сети Интернет произведений в сфере культуры, науки и образования	153
Установление основных категорий произведений культуры, науки и образования (аудио, видео, текст, изображение и др.) по форматам файлов и целям использования.....	153
Определение основных реестров произведений культуры, науки и образования (ISBN, DOI и др.).....	191
DOI как «система вещей» – претендующая на полноту онтология объектов и предметной деятельности.....	206
Раздел 4. Исследование использования системы DOI (Digital Object Identifiers) и построение пилотного проекта для работы с данной системой	217
Заключение	227