



- ▶ Журнал ЭБ
- ▶ О журнале
- ▶ Редакционная коллегия и редакция
- ▶ Что нового?
- ▶ 2014 год
- ▶ 2013 год
- ▶ 2012 год
- ▶ 2011 год
- ▶ 2010 год
- ▶ 2009 год
- ▶ 2008 год
- ▶ 2007 год
- ▶ 2006 год
- ▶ 2005 год
- ▶ 2004 год
- ▶ 2003 год
- ▶ 2002 год
- ▶ 2001 год
- ▶ 2000 год
- ▶ 1999 год
- ▶ 1998 год

▶ ENGLISH

## Электронные библиотеки -1999 - Том 2 - Выпуск 2

### Состояние работ по внедрению системы универсальных имен ресурсов Интернет для электронных информационных служб национальных библиотек

*В.М.Давыдов, Е.В.Давыдов, М.Е.Шварцман*  
*Российская государственная библиотека*

#### Введение

Постоянный рост объема информации, представленной в электронном виде и имеющей сетевую структуру, вызывает необходимость иметь надежные и универсальные способы наименования информационных ресурсов. Для этого организацией, занимающейся решением технических задач сети Интернет (IETF - Internet Engineering Task Force), принят и рекомендован к использованию термин URN - "Универсальное имя ресурса", т.е. имя, которое идентифицирует ресурс и не зависит от его расположения в Сети. По своему назначению URN должны быть глобально уникальны, постоянны и доступны при работе в Интернет.

Концепция универсальных имен в принципе с пониманием поддержана как специалистами по сетям, так и специалистами библиотечного дела, но до сих пор нет единого мнения по вопросам внедрения URN в практику. Однако в последнее время членами ведущих исследовательских групп ряда университетов и национальных библиотек [1-5], которые активно работают в этой области, достигнуты соглашения по большинству главных вопросов. Результаты этих соглашений сформулированы в виде стандартов и меморандумов Интернет [6].

Представляется, что в недалеком будущем система URN станет основой для идентификации электронных ресурсов, получения их библиографических данных, авторских "первоисточников" и их копий в глобальной сети Интернет.

С осени 1998 г. произошел значительный прогресс в работе по стандартизации системы URN. Рабочей группой URN организации IETF была завершена подготовительная работа, и новые проекты стандарта переданы в IETF для формального утверждения. Общее согласие в рассматриваемых вопросах весьма важно, тем более что до сих пор большинство специалистов, управляющих большими совокупностями интерактивной (on-line) информации, отказывались от использования какой-либо формы URN. Достигнутые соглашения имеют два главных результата:

- Пользователи, которые хотят дать постоянные имена интерактивным ресурсам, могут теперь планировать включение URN в документы, индексы и интерактивные системы. При этом они с большой вероятностью могут быть уверены, что будущие изменения форматов URN не будут вынуждать их переформатировать или изменять существующие URN.
- Использование согласованных стандартов позволит исключить неприятности, связанные с необходимостью изменять специфические (частные) схемы имен, обеспечит долговечность и полноценность назначенных URN, а также обеспечит непрерывную поддержку для существующих URN через другие системы разрешения имен в случае, если какая-либо схема имени перестанет поддерживаться в первоначальной форме. Таким образом, пользователи, которые назначают имена внутри любой из согласованных схем, могут быть уверены, что система не устареет.

Действенность существующих рекомендаций и стандартов по системе URN состоит в том, что ими можно руководствоваться при различных собственных подходах к решению данной проблемы. Кроме того, система имеет перспективу долгосрочного развития. Наименование ресурсов является комплексной проблемой, т.к. различные группы пользователей заинтересованы в URN по различным причинам. Они привносят и используют различные концепции и различные технические подходы. Поэтому достижение общего соглашения тем более полезно и необходимо для нахождения правильных технических решений. В целом работы по созданию и поддержанию глобальной и национальных систем URN можно разделить на три взаимосвязанные организационно-технические направления:

- регистрация "пространства имен",
- генерация URN и метаданных,
- "разрешение" информационных ресурсов на основе URN.

В настоящее время уже существуют несколько экспериментальных систем URN. Наибольших результатов в этой области достигли университетские и национальные библиотеки скандинавских стран - Финляндии, Швеции и Норвегии, однако по оценкам специалистов для завершения работ потребуется еще несколько лет.

## 1. Хронология международных совещаний по проблеме URN

В августе 1998 г. в Амстердаме состоялось совещание директоров национальных библиотек (CDNL - Conference of Directors of National Libraries), на котором было предложено провести исследования по вопросу использования URN как постоянных идентификаторов электронных документов в сети Интернет.

В октябре 1998 г. в Праге прошла конференция работников национальных европейских библиотек (CENL - Conference of European National Librarians), на которой рассматривались вопросы внедрения системы URN [2].

В январе 1999 г. Библиотека университета Хельсинки провела совещание по системе URN. В нем участвовали специалисты по ключевым системам идентификации электронных документов (URN, DOI - Digital Object Identifier, ISBN, ISSN).

В марте 1999 г. проведена встреча в рамках исследовательского проекта Computerised Bibliographic Record Actions (CoBRA+), осуществляемого при поддержке Европейской Комиссии.

## 2. Описание системы URN и планов ее развития

Хорошим введением в понятие URN является межсетевой меморандум RFC 1737 "Функциональные требования для универсальных имен ресурсов", действующий с 1994 г. URN позволяет идентифицировать документы (ресурсы) сети Интернет. Идентификатор URN является постоянным и уникальным; присвоенный документу URN он должен быть неизменным, если интеллектуальное содержание документа не изменилось. "Использованный" URN не может быть присвоен другому документу. Система "разрешения" должна давать пользователю возможность преобразовывать URN в URL или метаданные искомого документа.

URN имеют следующий синтаксис написания: <URN> ::= "urn:" <NID> ":" <NSS> (фразы, заключенные в кавычки, являются обязательными) NID означает Namespace Identifier - идентификатор "пространства имен", и является кодом, указывающим на конкретную систему URN и облегчающим интерпретацию синтаксиса NSS. NSS (Namespace Specific String) собственно и является кодом, который однозначно и единственно идентифицирует документ. Предполагается, что URN для системы NBN всегда должен начинаться с кода страны и следующей за ним информации - например, для Финляндии используются буквы "fi", год (YY) и число, увеличиваемое по мере присвоения URN. В результате примером записи URN для Финляндии может быть **URN:NBN:fi-fe976238**, а для Швеции - **URN:NBN:se-d199811234**.

Существует экспериментальная система для автоматической генерации URN на базе NBN для всех скандинавских стран (<http://www.lub.lu.se/dc/urntest.pl>). При присвоении URN применяются общие для всех соглашения:

1. URN может быть присвоен только документу, который является относительно стабильным и имеет значительное информационное содержание.
2. Копии документа должны использовать один и тот же URN, т.е. документ имеет один URN, но несколько URL.
3. Варианты ресурса в различных форматах файлов (например, текстовый файл в формате MS Word или HTML) должны все иметь один и тот же URN. Для того, чтобы улучшить процесс поиска, URN должен быть вставлен в метаданные о происхождении документа.
4. Если документ претерпел редакторские изменения, новый URN может быть присвоен только в случае изменения интеллектуального содержания.
5. Если размещаемый в сети документ имеет ISBN, ISSN или иной идентификатор ID, он может быть использован в качестве URN. Хотя NID для традиционных систем идентификации еще официально не определены, весьма вероятно, что для ISBN это будет "ISBN", для ISSN - "ISSN" и т.д. Например, URN:ISBN: <ISBN-number>, URN:ISSN: <ISSN-number>.
6. Документ должен иметь только один URN, но если документ состоит из нескольких отдельных файлов, они должны быть идентифицированы по отдельности.

После того, как URN получен, он должен быть внесен в текст идентифицируемого документа, а точнее - в раздел метаданных (META tag of HTML document). Только в этом случае системы индексации Интернет смогут обнаружить URN, и это приведет к желаемому результату, хотя в настоящее время системы типа Alta Vista Nordic или Web Index пока не способны использовать URN. Наилучшим способом внесения URN в HTML документ является использование элементов метаданных Dublin Core (см., например, [http://purl.oclc.org/metadata/dublin\\_core/](http://purl.oclc.org/metadata/dublin_core/)). В этом случае синтаксис написания строки метаданных имеет следующий вид: <META NAME="DC.Identifier" SCHEME="URN" CONTENT="URN code">

Например:

```
<META NAME="DC.Identifier" SCHEME="URN" CONTENT="URN:NBN:se-d199811234"> или  
<META NAME="DC.Identifier" SCHEME="URN" CONTENT="URN:NBN:fi-fe976242">
```

Если есть несколько версий одного и того же документа в разных форматах, то URN должен быть включен в каждый из них, причем для удобства нахождения первоисточника код URN должен быть внесен в соответствующую строку метаданных. Например:

```
<META NAME="DC.Source.Identifier" SCHEME="URN" CONTENT="URN:NBN:se-d199811234">
```

Написание "вручную" строк метаданных Dublin Core (DC) является довольно сложной задачей. Чтобы упростить процесс написания, разработаны специальные программы с заполнением форм по образцам и последующим автоматическим получением строк DC метаданных. В Интернет эти формы доступны, например, по адресу <http://www.lub.lu.se/cgi-bin/nmdc.pl> или <http://linnea.helsinki.fi/meta>. Эта программа разработана в рамках проекта скандинавских стран - Nordic Metadata Project и, кроме того, исходный текст программы был передан для использования в Национальную библиотеку Австралии. Аналогичная программа "генератор DC метаданных" доступна также на сервере Российской государственной библиотеки по адресу <http://www.rsl.ru/DC/>

К сожалению, инфраструктура "разрешения имен" на базе URN пока не создана, но в будущем автоматические поисковые системы, которые условно можно назвать "национальные серверы регистрации URN", будут собирать всю информацию о DC метаданных из зарегистрированных документов, следить за ее обновлением и предоставлять пользователям Интернет возможность получать "первоисточник" документа, его описание - библиографические данные, а также список адресов (URL), на которых размещены копии документа.

Система "разрешения" имен будет представлять собой двухступенчатый процесс. Глобальный регистр идентификации "пространства имен" будет использоваться для нахождения сервера - национального, международного или любого другого, содержащего информацию о зарегистрированных URN и отвечающего за информацию о соответствующем NID. Этому серверу будет передаваться запрос на получение документа. Например, если запрашиваемый NID будет "ISSN", то запрос будет передан серверу "разрешения" URN, поддерживаемому Центром ISSN, поскольку именно на этом сервере будет расположена действительная информация об URL электронных журналов, которые пользователь хотел бы найти.

По мере создания системы URN должны быть также соответствующим образом доработаны компьютерные программы типа Microsoft Internet Explorer и Netscape Navigator для того, чтобы пользователь мог получать документ по URN, а не только по URL, как это происходит до сих пор.

### **3. Связь и аналогия между URN и известными системами классификации**

Весьма важно оценить, как традиционные системы типа ISBN, ISSN или NBN вписываются в общую архитектуру системы универсальных имен ресурсов URN. Первым шагом для использования существующих идентификаторов в качестве URN должно быть определение соответствующих "пространств имен" для каждого идентификатора. Например, Международный центр ISSN должен зарегистрировать "пространство имен" - "ISSN", и только после того, как этот вопрос будет согласован, идентификаторы ISSN может официально использоваться в качестве URN. Следует подчеркнуть, что на совещании в Хельсинки представители международных центров ISBN и ISSN дали свое согласие, но заявили, что проведут работы по регистрации "пространства имен" URN для своих систем тогда, когда процедуры для выполнения этого будут полностью установлены.

Регистрация "пространства имен" для ISBN, ISSN и NBN позволит в ближайшем будущем использовать эти системы как URN. Это не вызовет дополнительной работы для национальных агентств ISBN или ISSN - за исключением того, что они должны будут сообщить издателям, как использовать, например, ISBN в качестве URN. Поскольку система URN не требует получения никакого дополнительного издательского идентификатора, то фактически URN окажется простым расширением к традиционному идентификатору. Все, что требуется - добавление **urn:isbn:** перед строкой ISBN. Заметим, что на возможность использования существующих библиографических идентификаторов ISBN, ISSN и SICI в качестве URN прямо указывает действующий документ RFC 2288.

Следующим шагом после регистрации "пространства имен" является создание глобальной компьютерной системы разрешения URN для зарегистрированных идентификаторов. При этом должно учитываться, что создание системы разрешения URN является не только технической, но также и до некоторой степени политической проблемой. Необходимо разработать систему, которая является в равной степени подходящей как для библиотек, так и для издателей.

Предварительное обсуждение на тему, как "разрешать" URN, основанные на ISBN, ISSN и NBN, содержится в [4]. "Разрешение" ISSN технически может оказаться очень простым как в случае "разрешения" библиографической информации, так и в том случае, когда цель состоит в фактическом выборе документа. Глобальная база данных ISSN, поддерживаемая Международным центром ISSN (<http://www.issn.org/>), может выдавать либо полное библиографическое описание или только список URL в основном всех электронных журналов, которые имеют URN на основе ISSN.

Что касается NBN, то, к сожалению, нет никакой единой "точки доступа" в Интернет, через которую можно было бы "разрешать" все URN на базе NBN. Эти услуги нужно обеспечивать на национальном

уровне. В каждой стране должен иметься один или несколько "сайтов", которые обеспечат услуги "разрешения имен". Так как номера NBN по определению не даются коммерческим документам, то по-видимому наиболее целесообразным может оказаться обеспечение услуги "разрешения имен" через службы, управляемые национальными библиотеками.

### Заключение

В данной статье обобщены результаты работ по созданию системы URN и рассмотрены рекомендации и стандарты, многие из которых все еще находящиеся в стадии согласования. Авторы выражают свою надежду, что система "Универсальных имен ресурсов" получит свое развитие и применительно к российскому сегменту Интернет.

### Ссылки

1. *Uniform Resource Names*, A Progress Report, D-Lib Magazine, February 1996, ISSN 1082-9873.
2. Esko Hakli, Juha Hakala (Helsinki University Library), *URN Implementation In National Libraries*, Discussion Paper, CENL, Prague, October, 1998 ([renki.lib.helsinki.fi/urn/urnimp.html](http://renki.lib.helsinki.fi/urn/urnimp.html)).
3. Juha Hakala, Traugott Koch (Finnish and Swedish National Libraries), *URN creation tool, Nordic Metadata Project*, March 1998 ([www.ub2.lu.se/metadata/URN-help.html](http://www.ub2.lu.se/metadata/URN-help.html), [linnea.helsinki.fi/meta/URN-help.html](http://linnea.helsinki.fi/meta/URN-help.html)).
4. Juha Hakala, *URN/DOI Initiative: A State of the Art - report*. Discussion paper for the CoBRA+ meeting, Helsinki University Library, 19.3.1999
5. R. Daniel (Los Alamos National Laboratory), *A Modest Suggestion For The URN->URL Service*, ([www.acl.lanl.gov/URI/initial\\_report.txt](http://www.acl.lanl.gov/URI/initial_report.txt)).
6. Документы Network Working Group (Request for Comments -RFC):
  - **RFC 1630**, T. Berners-Lee (CERN), *Universal Resource Identifiers in WWW, A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network as used in the World-Wide Web*, June 1994.
  - **RFC 1737**, K. Sollins (MIT/LCS), L. Masinter (Xerox Corporation) *Functional Requirements for Uniform Resource Names*, December 1994.
  - **RFC 2141**, R. Moats (AT&T), *URN Syntax*, May 1997.
  - **RFC 2168**, R. Daniel (Los Alamos National Laboratory), M. Mealling (Network Solutions, Inc.) *Resolution of Uniform Resource Identifiers using the Domain Name System*, June 1997.
  - **RFC 2169**, R. Daniel (Los Alamos National Laboratory), *A Trivial Convention for using HTTP in URN Resolution*, June 1997.
  - **RFC 2288**, C. Lynch (Coalition for Networked Information), C. Preston (Preston & Lynch), R. Daniel (Los Alamos National Laboratory), *Using Existing Bibliographic Identifiers as Uniform Resource Names*, February 1998.

---

© Давыдов В.М., Давыдов Е.В., Шварцман М.Е., 1999

Последнее обновление страницы было произведено: 2003-12-09

Все предложения и пожелания по содержанию и структуре портала направляйте по адресу [rdlp@iis.ru](mailto:rdlp@iis.ru)

