

В. Б. Барахнин¹, А. М. Федотов²

¹Институт вычислительных систем СО РАН
пр. Акад. Лаврентьева, 6, Новосибирск, 630090, Россия
E-mail: bar@ict.nsc.ru

²Новосибирский государственный университет
ул. Пирогова, 2, Новосибирск, 630090, Россия
E-mail: fedotov@nsu.ru

ИНФОРМАЦИОННАЯ СИСТЕМА: ВЗГЛЯД НА ПОНЯТИЕ¹

Введение

В свое время Альберт Эйнштейн предупреждал, что до конца двадцатого века человечество станет свидетелем «информационного взрыва». Сегодня можно утверждать, что этот «взрыв» произошел, причем сила и скорость его «ударной волны» растут с каждым днем. Причина его возникновения связана не столько с гигантским прогрессом в информационных технологиях, сколько с возросшим во много раз потоком информации, необходимой для жизни современного общества.

Проблема доступа к информации является одной из основных проблем, возникающих в деятельности научного исследователя. Любой научный процесс порождает огромные объемы данных, и работать с ними становится все сложнее по мере того, как гигабайты данных превращаются в терабайты. Количество данных может когда-нибудь превысить способность компьютеров их обрабатывать, поэтому необходимы новые инструментальные средства и алгоритмы для анализа этих данных. Вместе с тем предъявляются серьезные требования к обеспечению прозрачного доступа и долговременной сохранности информации. В результате вопросы «что хранить?» и «как хранить?» остаются самыми существенными: без ответа на них все остальные проблемы работы с информацией теряют актуальность.

На первый взгляд может сложиться впечатление, что развитие информационных технологий уже само по себе способно вывести работу с научной информацией на качественно новый уровень, но, к сожалению, это совсем не так. Современные вычислительная техника и информационные технологии предоставляют исследователю мощный аппарат для манипулирования с *данными*, но не с *информацией*. Данные представляют информационную ценность лишь тогда, когда они являются составной частью некоторой модели реального мира и связаны с другими данными [Колмогоров, 1965; 1987]. Как отмечал А. А. Ляпунов [1980], «информация всегда относительна, она зависит... от того, какой информационной системой она воспринимается». Иными словами, конечная цель работы, связанной с применением информационных технологий – *понимание* того или иного явления, а не получение каких-либо чисел или картинок.

Наше представление о явлениях и объектах окружающего мира складывается из обобщения данных, получаемых в процессе наблюдений или экспериментов при описании отдельных элементов того или иного объекта. Поэтому важнейшей задачей является построение *информационной модели* объекта, которая связывает разнообразные данные, превращая тем самым их в информацию. Здесь необходима разработка иерархической системы понятий, определения отношений общих и частных понятий, определения характера связей между ними, разработка стандартов описания понятий и форматов данных.

Для наиболее продуктивной работы исследователю нужны данные («информация»), представленные в виде *знаний* – «адекватного отражения действительности в сознании человека в виде представлений, понятии, суждений теорий» [Философский..., 1983]. Для представления «модели данных», отражающей свойства реального мира, используется понятие онтологии, понимаемая в данном случае как набор определений (на формальном языке) фрагмента декларативных знаний, включающий термины, типы и соотношения (аксиомы), описывающие фрагмент знания [Нариньяни, 2001]. Такое понимание онтологии отражает возможность перехода от конструирования ограниченных «разовых» моделей предметных областей к описаниям знаний о предметных областях, способных становиться частями более общих моделей знаний.

Для обработки данных, структурированных при помощи той или иной формальной модели, применяются комплексы аппаратно-программных средств, обычно называемые *информационными системами*. В настоящее время среди широкого класса информационных систем наибольшее распространение получили *документально-фактографические информационно-поисковые системы* [Словарь по кибернетике, 1989], осуществляющие поиск, переработку и хранение как документальной, так и фактографической информации (под последней фактически понимаются знания), причем в этих системах, как правило, предполагается наличие простейших видов логического и эвристического вывода.

Отметим одно важное ограничение, связанное со сложностью рассматриваемых систем. Под общее определение информационно-поисковой системы формально подпадают и такие комплексы аппаратно-программных средств, в которых реализованы лишь простейшие поисковые запросы типа выдачи документа по его известному имени. Однако выведение подобных комплексов за рамки информационно-поисковых систем сделано еще в 1960-е гг. в монографии [Михайлов и др., 1968], где подчеркнута, что устройства и машины, предназначенные лишь для отыскания документов по известным адресам их хранения, информационно-поисковыми системами не являются. Такой подход соответствует фундаментальным положениям работы [Ляпунов, Яблонский, 1963], согласно которой характерной чертой управляющих (в широком понимании этого термина) систем, требующих специальных кибернетических рассмотрений, является их сложность, проявляющаяся в большом количестве элементов, сложной системе связей, больших потоках информации и реализации сложных функций. Подчеркивание этой же особенности информационных систем характерно и для англоязычной терминологии [Webster's New..., 1992]. Далее вместо термина «документально-фактографические информационно-поисковые системы» там, где это не вызовет недоразумения, мы будем для краткости использовать термин «информационные системы».

Основная проблема, связанная с построением теории информационных систем, – отсутствие общей теории, связывающей воедино *данные, информацию и знания*. Имеется множество подходов к понятию информация с философских или кибернетических позиций, включая так называемую «техническую» теорию информации, которую следовало бы называть «теорией передачи и хранения данных». Можно обнаружить десятки и сотни порой противоречащих друг другу определений того, что является информацией или знанием.

Заметим, что работы по автоматическому получению знаний в рамках моделей искусственного интеллекта (см., например: [Гаврилова, Хорошевский, 2000]) на базе так называемых [Словарь по кибернетике, 1989] информационно-логических систем нацелены на решение узкоспециализированных задач, причем модель знаний строится на основании учета мнений целого коллектива экспертов, поскольку отдельно взятый специалист (еще раз подчеркнем – специалист в узкой области!) не обладает достаточными знаниями для решения подобных задач.

Опираясь на интуицию, эксперты приходят к выводу о порочности нынешней ситуации, но о каком-либо серьезном переосмыслении проблем данных и контента пока речь не идет. Показательно высказывание Сьюзан Фельдман, ведущего аналитика IDC, специализирующейся на технологиях работы с «контентом»: «Сегодня весь поисковый рынок устремился в область конвергенции „содержания“ с данными, поскольку, в конечном счете, ценность имеет информация» [Черняк, 2005].

Исходя из накопленного нами опыта создания распределенной информационной системы СО РАН, попробуем сформулировать наиболее общие принципы построения моделей информационных систем, работающих как со структурированными данными («информацией»), так и со знаниями, и укажем отличительные особенности таких систем.

Информационные системы как объект исследований системного анализа

Проанализируем информационные системы с использованием методики общей теории систем. Отметим, что классическое определение системы как «множество объектов вместе с отношениями между объектами и между их атрибутами» [Холл, 1969] основано на тех же понятиях, что и, например, реляционная модель данных [Codd, 1970].

Основные системные принципы включают в себя [Философский..., 1983]:

- целостность (зависимость каждого элемента, свойства и отношения от его места и функций внутри целого);
- структурность (возможность описания системы через установление ее структуры, т. е. сетей связей и отношений системы);
- иерархичность (каждый компонент системы в свою очередь может рассматриваться как система, а исследуемая система – как компонент более широкой системы);
- множественность описания (посредством использования множества различных моделей);
- взаимозависимость системы и среды (система формирует и проявляет свои свойства в процессе взаимодействия со средой, являясь при этом активным компонентом взаимодействия).

Применительно к кибернетическим системам методология системного анализа была описана в статье [Ляпунов, Яблонский, 1963]. Выделены два основных подхода к изучению кибернетических систем: макроподход, при котором система рассматривается как «черный ящик» для исследования ее взаимодействия с окружающей средой, и микроподход, при котором изучается внутреннее строение системы. В рамках этих подходов сформулированы двенадцать основных направлений исследования систем (в рамках макроподхода – информационные потоки, коды, функции, функционирование систем; в рамках микроподхода – элементы, связи между элементами, алгоритмизация, анализ, синтез, преобразования, эволюция, надежность систем). При этом подчеркнута, что некоторые из перечисленных направлений (прежде всего, преобразования и эволюция управляющих систем) актуальны далеко не для всех типов систем, включая лингвистические (к которым относятся и информационные системы).

Нетрудно видеть, что в рамках макроподхода изучаются вопросы взаимозависимости системы и среды, в то время как комплекс проблем, изучаемых в рамках микроподхода, фактически приводит к исследованию целостности системы. Рассматривая отдельные аспекты микроподхода, можно отметить, что изучение элементов системы (в том числе выявление масштаба основных элементов-«кирпичей») требует исследования ее иерархичности, изучение связей – исследования структурности, анализ системы требует использования множества различных моделей ее описания.

Рассмотрим подробнее, как при создании информационных систем реализуются основные системные принципы.

Вопросы макроподхода (прежде всего, взаимозависимости информационной системы и среды) могут представлять интерес как в плане взаимодействия с пользователем – разработка моделей информационных запросов и моделей представления информации, так и в плане отражения системой изменений во внешней среде (т.е. актуализации информации), а также интероперабельности – интеграции системы внешними информационными системами.

Актуализация информации является слабым местом практически всех информационных систем некоммерческой направленности (за исключением, разумеется, систем, поддерживаемых государственными органами), предназначенных для функционирования в течение неопределенно долгого времени. Причина этого очевидна – недостаток средств, прежде всего, для оплаты труда лиц, которые должны отслеживать изменения информации. Это приводит к необходимости максимальной автоматизации процесса актуализации информации. Некоторые

аспекты решения этой задачи рассмотрены в [Барахнин, Ведерников, 2006]. Одним из основных требований к информационной системе, процесс пополнения которой автоматизирован, является наличие в системе одного или нескольких каталогов. В противном случае, при отсутствии явно созданного множества унифицированных структурированных описаний документов, автоматическое добавление в систему новых документов становится, очевидно, крайне проблематичным.

Низкая интероперабельность – также весьма существенный недостаток, негативно отражающийся на важнейшей функции информационных систем: организации поиска документов по их атрибутам. Различают два уровня интероперабельности: семантический и технический [Фейгин, 2004], причем в последнем выделяют иногда синтаксический уровень [Бездушный и др., 2005]. Семантическая интероперабельность, заключающаяся в использовании согласованных стандартов метаданных (обзор которых применительно к научным информационным системам см. [Там же]), как правило, соблюдается. Проблемы возникают на уровне технической интероперабельности, точнее, согласования моделей данных и форматов их представления (что относится к синтаксической интероперабельности). Наиболее естественной формой унификации представления данных является каталог.

Переходя к микроподходу, важно подчеркнуть, что информационная система должна оперировать не непосредственно с документами (т. е. целостными информационными объектами, которые описывают, представляют, отображают или моделируют некоторые сущности реального мира), а с их поисковыми образами. При отсутствии поставленных в соответствие документам поисковых образов поиск документа возможен лишь по его адресу, что противоречит приведенному выше определению информационной системы.

Заметим, что в некоторых случаях содержание исходного документа может входить в поисковый образ в качестве одного из элементов (это противоречит ограничению из монографии [Михайлов и др., 1968], но из контекста данной работы следует, что подобное ограничение было вызвано необходимостью уменьшения объема поисковых образов с целью уменьшения трудоемкости процесса их обработки). С другой стороны, поисковый образ документа тоже является документом (описывающим исходный документ), поэтому далее, говоря об информационных системах, мы будем использовать термин «документ» в значении «поисковый образ исходного документа». Следует подчеркнуть, что в классических работах по информатике и кибернетике [Словарь по кибернетике, 1989; Михайлов и др., 1968], вышедшим, в том числе, в конце 1980-х гг., поисковый образ документа не рассматривается даже в качестве вторичного документа.

Итак, основной элемент-«кирпич» (в терминологии [Ляпунов, Яблонский, 1963]) информационной системы – документ, который представляют собой информационный ресурс, имеющий (по определению [Berners-Lee et al., 1998]) уникальный идентификатор и обладающий некоторой структурой и содержанием.

Документы в информационной системе описываются посредством метаданных, причем структура метаданных иерархична. Наиболее общий характер имеют метаданные, задающие структуру документа, т. е. описывающие метаданные более низкого уровня (атрибуты документа), которые определяют содержание документа. Наконец, значения этих атрибутов является фактически метаданными по отношению к исходному документу. Отсюда следует важнейшая отличительная черта информационной системы: она работает не с данными, а исключительно с метаданными.

Теперь мы можем сформулировать основные особенности реализации системных принципов для информационных систем.

Целостность системы проявляется в зависимости каждого объекта, свойства и отношения от его места и функций внутри целого и реализуется посредством использования единого набора метаданных $M = \cup M^i$. Тем самым любой документ d_i системы представляется как $d_i = \langle m_i^{j,k} \rangle$, где $m_i^{j,k}$ – значения элементов метаданных M^j , k – количество значений (с учетом повторений) соответствующего элемента метаданных в описании документа.

Иерархичность информационной системы проявляется в том, что она состоит из, вообще говоря, разнородных подсистем, отвечающих тем или иным частным задачам. Документы, описываемые при помощи одних и тех же элементов метаданных, образуя множество

$M_i \subseteq M$, образуют класс K_i . Если $M_1 \subset M$, $M_2 \subset M$ и $M_1 \subset M_2$, то класс K_2 является подклассом класса K_1 . Множество унифицированных структурированных документов-описаний одного класса называют *каталогом*. Фактически каталог объединяет поисковые образы исходных документов.

Структурность системы обеспечивается оптимальным выбором модели связей между документами, позволяющей адекватно описывать различные аспекты соответствующих межсущностных отношений. Достаточно универсальный характер имеет, например, модель направленных связей [Барахнин, Леонова, 2005]. Подробнее, если документ d_i входит в качестве значения элемента M^j метаданных документа d_j , то можно говорить о связи между этими документами вида $M^j < d_i, d_j, m_{i,i}^{j,k} >$, где $m_{i,i}^{j,k}$ – атрибуты этой связи, являющиеся значениями соответствующих элементов метаданных. Таким образом, выстраиваемые отношения фактически переносятся на уровень элементов, определяющих структуру документов.

Множественность описания системы подразумевает наличие множества различных аспектов построения системы (информационная модель системы, содержательное наполнение, используемые технологии и пр.). Наиболее общий характер имеет описание *модели информационной системы*, которая строится посредством задания классов K_i , определяемых соответствующими множествами элементов метаданных M_i , и типов возможных связей между классами $M^j < K_i, K_i >$ с указанием элементов метаданных $M_{i,i}^j$, описывающих атрибуты соответствующих связей. Таким образом, для построения модели информационной системы используется комбинация иерархической и реляционной моделей данных.

Конкретное наполнение информационной системы определяется содержанием ее каталогов, причем, поскольку иерархическая модель данных может быть представлена в реляционном виде, есть смысл говорить и о каталогах связей.

Модели поиска информации и извлечения знаний

Как уже отмечалось выше, важнейшей, с точки зрения пользователя, чертой информационных систем (и, более того, их определяющей характеристикой) является наличие возможности осуществлять сложные информационные запросы. Можно выделить три основные модели поиска в информационных системах:

- контекстный поиск;
- атрибутивный поиск;
- поиск по аналогии.

С точки зрения организации поиска атрибуты (поисковые признаки) удобно подразделять на три типа: текстовые; числовые; табличные (т. е. выбираемые из заданного списка).

Контекстный поиск отбирает документы, у которых значения атрибутов текстового типа (любых или выбранных пользователем) содержат заданный в качестве поискового запроса текстовый фрагмент. В некоторых случаях по желанию пользователя учитываются возможные морфологические изменения текстового фрагмента. Запрос к каждому атрибуту может также включать несколько текстовых фрагментов, связанных логическими операциями конъюнкции, дизъюнкции и отрицания. Этими же операциями могут быть связаны запросы к нескольким атрибутам.

Контекстный поиск – наиболее простая модель поиска, которая может быть реализована даже при минимальной структуризации документов, однако ее эффективность зачастую не велика. Она максимальна в том случае, когда в качестве поискового запроса пользователь задает заранее известную точную цитату достаточной длины. В противном случае, если поисковый запрос слишком короток, информационная система выдаст большое количество документов, многие из которых наверняка будут не интересны пользователю. Напротив, для запросов большой длины зачастую вообще не обнаруживается соответствующих им документов, поскольку авторы документов могли описывать интересующий пользователя предмет фразами, несколько отличающимися от заданной в запросе.

При атрибутивном поиске информационный запрос представляет собой набор значений одного или нескольких атрибутов, выбираемых из списка (и/или значений числовых атрибутов из заданного диапазона), связанных логическими операциями (обычно – конъюнкцией).

Атрибутивный поиск целесообразен в тех случаях, когда пользователя интересует не столько какой-либо конкретный документ, сколько класс документов, удовлетворяющий некоторому набору свойств.

Поиск по аналогии предполагает задание в качестве информационного запроса непустого множества документов. В качестве результата запроса выдаются документы, каждый из которых в определенном смысле близок к одному из документов, входящих в заданное множество.

Количественная характеристика меры сходства определяется на множестве документов D следующим образом: $m: D \times D \rightarrow [0, 1]$, причем функция m в случае полного сходства принимает значение 1, в случае полного различия – 0. Вычисление меры сходства осуществляется по формуле вида $m(d, D') = \sum a_i m_i(d, D')$, где i – номер атрибута, a_i – весовые коэффициенты, причем $\sum a_i = 1$, $m_i(d, D')$ – мера сходства по i -му атрибуту, вычисляемая как $m_i(d, D') = \max_j m_i(d, d_j)$, d_j – документы из множества D' . Для табличных атрибутов мера сходства по i -му атрибуту определяется следующим образом: если значение i -го атрибута данного документа принадлежит соответствующей проекции множества D' , то мера близости равна 1, иначе 0. В случае атрибута, принимающего множественные значения, $m_i = n_{0i}/n_i$, где n_i – общее количество элементов, составляющих значение i -го атрибута документа, n_{0i} – количество элементов, принадлежащих проекции D' .

Наконец, чтобы информационно-поисковая система могла быть использована для получения новых знаний, ее пользователю должна быть предоставлена возможность проверять, истинно или нет утверждение $R_S(d_1, \dots, d_n)$ относительно сущностей, описываемых документами d_1, \dots, d_n . Высказыванию $R_S(d_1, \dots, d_n)$ формально соответствует n -местный предикат P_S , определенный на множестве документов, причем при его построении могут использоваться определенные документы системы (точнее, значения атрибутов этих документов), информация из онтологии предметной области и т. п. Если количество документов в системе, способных выступать в качестве аргументов предиката P_S , велико, то весьма перспективным методом извлечения новых знаний является проверка истинности предиката на различных наборах документов, автоматически перебираемых системой. Подчеркнем отличие такого подхода от подхода, применяемого специалистами в области искусственного интеллекта для разработки экспертных систем: последние предназначены для решения узкоспециализированных задач, содержат относительно небольшой объем документов, и основной упор при их создании делается на развитие сложных продукционных правил. В нашем же случае правила вывода могут быть достаточно просты, а получение новых знаний возможно благодаря большому объему документов, способных выступать в качестве аргументов проверяемых утверждений.

Таким образом, для эффективной организации информационного поиска и получения новых знаний важна хорошая структуризация документов, предусматривающая, в частности, достаточно большое количество поисковых признаков, образующих поисковый образ документа. Это накладывает особые требования на каталог информационно-поисковой системы, которые мы сформулируем в следующем пункте.

Особенности организации каталога информационно-поисковой системы

Как отмечалось выше, каталог информационно-поисковой системы объединяет поисковые образы исходных документов. Возникает вопрос: какой набор поисковых признаков требуется выделить для того, чтобы выполнялось основное свойство информационно-поисковых систем – возможность реализации поисковых запросов более сложных, чем выдача документа по известному имени?

Покажем, что реальные возможности контекстного поиска фактически сводятся к поиску документов с известным именем (частью имени) или известной фразой из текста документа. Очевидно, эффективность контекстного поиска максимальна, если в качестве поискового запроса пользователь задает заранее известную точную цитату достаточной длины. В противном случае, если поисковый запрос слишком короток, информационная система выдаст множество документов, большинство из которых наверняка будут не пертинентными (т. е. не соответствующими информационной потребности пользователя), хотя и формально релевантными (т. е. соответствующими «букве» информационного запроса). Для «произволь-

ных» запросов большой длины зачастую вообще не обнаруживается даже релевантных документов, поскольку авторы документов могли описывать интересующий пользователя предмет фразами, несколько отличающимися от заданной в запросе.

Описываемая ситуация, когда система позволяет осуществлять только контекстный поиск, возникает, если информационно-поисковый язык системы практически совпадает с естественным языком.

Более того, аналогичные проблемы встают и тогда, когда в информационно-поисковом языке отсутствуют средства выражения имманентных отношений между предметами, т. е. язык не имеет парадигматических отношений (примером такого языка может служить система унитаров – набора одиночных ключевых слов (в редких случаях словосочетаний)).

Таким образом, возможность получения в результате поискового запроса пертинентных документов появляется лишь в том случае, когда информационно-поисковый язык имеет средства выражения имманентных отношений, т. е. обладает словарем-тезаурусом. Так как определение тезауруса [Словарь по кибернетике, 1989] включает наличие смысловой классификации терминов, то наличие тезауруса означает возможность осуществления атрибутивного поиска (а также поиска по аналогии). В качестве поисковых атрибутов могут выступать разделы соответствующего классификатора. Итак, наличие тезауруса (и, следовательно, классификатора) в качестве составной части информационно-поискового языка, используемого при создании каталога, является необходимым и достаточным условием выполнения основного требования к информационно-поисковым системам – возможности реализации сложных информационных запросов.

Заключение

Резюмируя сказанное выше, под «информационной системой» мы будем понимать только такие комплексы аппаратно-программных средств для работы со структурированными данными, которые позволяют осуществлять информационный поиск документов не только по их именам, но и по атрибутам. Важнейшая отличительная черта информационной системы состоит в том, что: она работает не с данными, а исключительно с метаданными, при этом информационная система обязательно снабжена каталогом – множеством унифицированных структурированных документов-описаний (фактически объединяющем поисковые образы исходных документов).

Для эффективной организации информационного поиска важна хорошая структуризация документов, обеспечиваемая выбором адекватной модели информационной системы, которая позволяет выделить достаточное количество атрибутов документа, выступающих в качестве поисковых признаков, образующих поисковый образ документа. При этом для организации сложных информационных запросов необходимо и достаточно, чтобы информационно-поисковый язык, при помощи которого описывается каталог системы, обладал тезаурусом.

Список литературы

Барахнин В. Б., Ведерников В. В. Автоматизированная каталогизация электронных журнальных публикаций // Тр. междунар. конф. «Вычислительные и информационные технологии в науке, технике и образовании». Павлодар, 20–22 сентября 2006 г. Т. 1. С. 209–214.

Барахнин В. Б., Леонова Ю. В. Информационная модель отношений между документами в информационной системе // Выч. технологии. 2005. Т. 10. Спец. вып. С. 129–137.

Бездушный А. Н., Кулагин М. В., Серебряков А. А. и др. Предложения по наборам метаданных для научных информационных ресурсов // Выч. технологии. 2005. Т. 10. Спец. вып. С. 29–48.

Гаврилова Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. СПб.: Питер, 2000.

Колмогоров А. Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации. 1965. Т. I, вып.1. С. 3–11.

Колмогоров А. Н. Теория информации и теория алгоритмов. М.: Наука, 1987.

Ляпунов А. А. О соотношении понятий материя, энергия и информация // В кн.: А. А. Ляпунов. Проблемы теоретической и прикладной кибернетики. Новосибирск: Наука, 1980. С. 320–323.

Ляпунов А. А., Яблонский С. В. Теоретические проблемы кибернетики // Проблемы кибернетики. 1963. Вып. 9. С. 5–22.

Михайлов А. И., Черный А. И., Гиляревский Р. С. Основы информатики. М.: Наука, 1968.

Нариньяни А. С. Кентавр по имени ТЕОН: Тезаурус + Онтология // Труды междунар. семинара «Диалог'2001 по компьютерной лингвистике и ее приложениям». Аксаково, 2001. Т. 1. С. 184–188.

Словарь по кибернетике. Киев: Главная редакция УСЭ им. М. П. Бажана, 1989.

Фейгин Д. Концепция SOA // Открытые системы. 2004. № 6. [http://www.osp.ru/os/2004/06/184447/_p1.html].

Философский энциклопедический словарь. М.: Советская энциклопедия, 1983.

Холл А. Д., Фейджин Р. Е. Определение понятия системы // Исследования по общей теории систем. М.: Прогресс, 1969. С. 252–282.

Черняк Л. От информационно-поисковых систем к корпоративному поиску // Открытые системы. 2005, № 11 [<http://www.osp.ru/os/2005/11/380532/>].

Berners-Lee T., Fielding R., Masinter L. Uniform Resource Identifiers (URI). Generic Syntax // RFC 2396. 1999. [<http://www.ietf.org/rfc/rfc2396.txt>].

Codd E. F. A relational model of data for large shared data banks // Comm. ACM 13,6 (June 1970). P. 377–387.

Webster's New World Dictionary of Computer Terms: 4th ed. N. Y.: Prentice Hall, 1992.

Материал поступил в редколлегию 3.09.2007