

Перспективы Интернета

Всегда онлайн

Шире, быстрее, мобильнее. До сегодняшнего дня это были три главные цели Интернета — даже тогда, когда мы не могли найти в Сети то, что нам было необходимо. Решение — сеть, которая понимает нас и саму себя.

Лучший друг тот, кто всегда там, где он нужен, кто знает наши интересы и нужды, кто вовремя скажет, что пора работать или где можно наконец-то купить диск с фильмом «Звездные войны. Эпизод 2». Есть ли такой друг из плоти и крови, зависит от вас, а вот в цифровой форме его можно будет скоро увидеть. Речь идет о так называемом «цифровом друге», разработанном Фраунхофским институтом программного обеспечения и системной техники (ISST), системе, которая круглые сутки снабжает вас необходимой информацией из Интернета.

Однако есть одно условие для существования такого сервиса: поисковые машины наконец-то должны начать понимать, что ожидает увидеть музыкальный фанат, когда вводит слово «Мадонна». И зачем ему нужна самая быстрая DSL-связь, если она выдает замороженных Мадонн или порнокартинки?

Действительно удобным для пользователя будет, если он сообщит «цифровому другу» всю информацию о своих привычках и интересах. Этот «другок» узнает место пребывания пользователя на данный момент по сети мобильной телефонии, из ка- »

» лendarя, организера или благодаря «умному дому». В случае необходимости он возникает в мобильном телефоне, в автомобильном компьютере или на дисплее микроволновой печи.

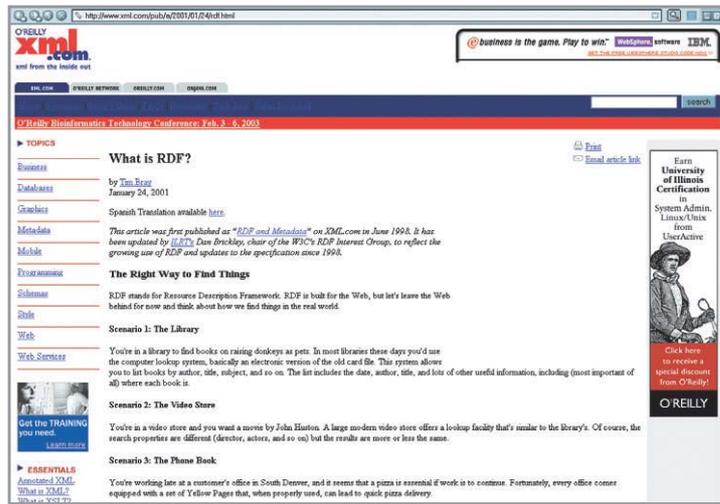
Фантазия, скажете вы? Напротив. «Уже до 2005 года этот независимый цифровой помощник станет реальностью», — убежден Карстен Линеманн (Carsten Linemann), разработчик ISST.

Уже сегодня система поддерживает каналы издания SMS, Fax, WAP, E-mail и HTML. «Однако полностью законченной система станет лишь тогда, когда компьютер будет в состоянии анализировать все данные в Сети», — говорит Тим Бернерс-Ли. Создателям Интернета представляется такая умная сеть, где сайты станут понятными для машин.

Сеть учится понимать свои данные

Как только это произойдет, придет третье поколение Интернета. Необходимость в этой, как ее называет Бернерс-Ли, «семантической сети» возникла уже давно — с тех пор, как поисковые машины стали подобны библиотекарям, снующим между полками, и при наличии такой массы данных не могут находить необходимое.

Обычная проблема: поисковые машины думают не ассоциативно. Они не знают, что подразумевается под «банком» — здание, где хранятся деньги, или карточ-



« На сайте www.xml.com уже есть информация по языку RDF

ный термин. Поэтому фундамент умной сети в том, чтобы создавать web-сайты так, чтобы компьютер был в состоянии понимать их содержимое.

С технической точки зрения здесь необходимы три компонента: гибкий язык программирования для передачи содержания, метод для понимания содержания, то есть правила, по которым можно автоматически связать содержание и значение. Таким многогранным метаязыком является eXtensible Markup Language, сокращенно XML. Программное обеспечение, которое может обращаться с сайтами, созданными на XML, способно, например, различить в адресе индекс и автоматически подсчитать расстояние до этого места. Для того чтобы наряду со структурой было

охвачено и значение содержания, необходим Resource Description Framework (RDF). RDF служит для того, чтобы выразить соотношение между двумя содержаниями, например одного сайта к другому. К тому же система различает значения одного и того же слова, как в описанном выше случае с «банком».

Третий компонент необходим для того, чтобы компьютер мог однозначно организовать содержание сайта и его значения. Документы и классификации, которые исполняют эти задания, называются онтологиями. Так, каждый знает, что WWW означает то же самое, что и World Wide Web, и принадлежит Интернету.

Эта система сравнима с методом в биологии делить группы с одинаковыми при- »

ТРИ ЭПОХИ ИНТЕРНЕТА



Час рождения Интернета — ученые американского Министерства обороны впервые соединили между собой четыре компьютера. Отличительной чертой явилось то, что компьютеры находились на большом расстоянии друг от друга

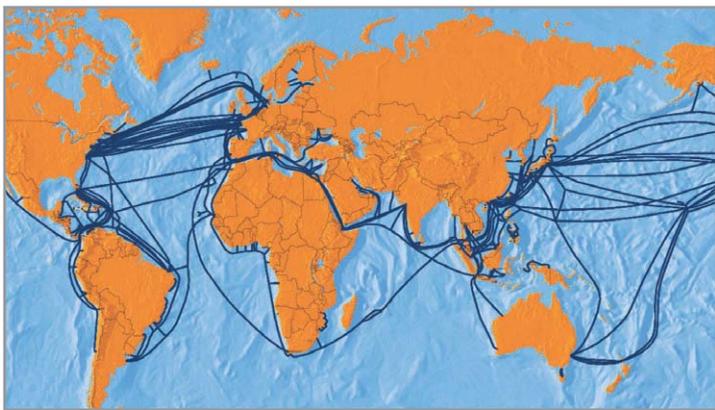
E-mail — Рей Томлинсон (Ray Tomlinson) переслал первое письмо, используя в адресе в качестве разделительного значка @

TCP/IP — Боб Кан (Bob Kahn) и Вinton Серф (Vinton Cerf) разработали единый формат для передачи данных и установления связи в Сети

WWW — Тим Бернерс-Ли (Tim Berners-Lee) и Роберт Келло (Robert Caillau) разработали HTML и создали первый сайт

Браузер — Марк Андерсен (Marc Andersen) при помощи Mosaic создает первый web-браузер, после которого появился Netscape Navigator

Internet 1 1968 1971 1974 1990 1992



◀ Подводные кабели уже оплели всю планету

» знаками на виды и подвиды. Первые сайты, которые были созданы по системе онтологии RDF, существуют уже два года. Их количество очень мало, так как скопление онтологий, как на Yahoo, закладывается вручную. «Для осуществления автоматической закладки, — говорит Гарольд Боули из Центра компетенции семантической сети в Кайзерслаутерне, — необходима еще пара лет. А затем она будет существовать только в экспертных системах с ограниченным количеством тем».

Прорыв такого спецприменения Боули ожидает через два года. Японские разработчики в компании Sony работают над «Person Memory», с помощью которой в семейных интернет-альбомах фотографию можно будет искать не только по имени файла, но и по дате или месту снимка. Необходимые данные уже сегодня включа-

ются в изображения, снимаемыми камерами Digicam.

Еще одно преимущество семантической сети: наряду с текстовой информацией можно найти и мультимедийное содержание. В недрах Сети до сих пор теряются картинки, музыкальные произведения или фильмы, если слово для поиска не стоит в web-адресе или названии файла. Чтобы это изменить, MPEG (Moving Pictures Expert Group) в прошлом году принял MPEG-7 как стандарт. Язык описания для аудио и видео должен позволить копаться в аудиовизуальном контенте и искать по ключевым словам, свойствам или общим чертам. Особенно хорошо переносится принцип семантической сети с его расширениями MPEG-7 на аудиоданные — благодаря Фраунхофскому институту интегрированных схем (IIS) и автоматической

идентификации аудиоданных, сокращенно AudioID. Если этой системе проиграть музыкальное произведение с CD или как файл MP3, она узнает песню через несколько секунд.

Система разбивает произведение на спектральные компоненты, подсчитывает признаки, такие как тональность и шумы, и сохраняет их наряду с названием и исполнителем. При банке данных в 100 000 названий процент того, что вы найдете необходимое, равен 99,5%. С такой точностью AudioID могла бы существенно изменить лицо музыкального рынка в Сети: для защиты авторского права или для более быстрого поиска на музыкальных порталах. В общем, эта технология подходит также и для фонограмм фильмов и телепередач, но для этого она должна быть немного модифицирована.

Сейчас IIS ищет финансовых партнеров, которые могли бы превратить эту систему в ходовой товар. На данный момент США и Европейский союз проверяют музыкальную индустрию AudioID и аналогичные технологии.

Как утверждает Юрген Герге — разработчик в IIS, — выводы об одном из методов будут известны уже в этом году. Между тем IIS сам пытается интегрировать AudioID в маленькие приборы. Первые опыты показали, что песни, транслируемые по радио, можно передавать по мо-



```

<!--></SCRIPT></noscript
src="http://chip.iwebb
bin/iww/CP/Nonepage;14/4/100
r=" width="1" height="1">
src="http://77
pic.gif?n=chipac=31160-3160
&tc=881&id=2&tp=&random=58
  
```

Хитрая сеть — первые сайты поддаются машинальному анализу

MPEG 7 — данный стандарт позволяет совершать аудиовизуальный поиск

Internet 3

2001

Internet 2 1998

1999

2000

Высокая скорость — В США и Германии разработчики с помощью независимых оптико-волоконных сетей стерли границы Интернета

WAP/I-MODE — Интернет становится мобильным — оба протокола мобильной связи оснащают телефоны web-браузерами

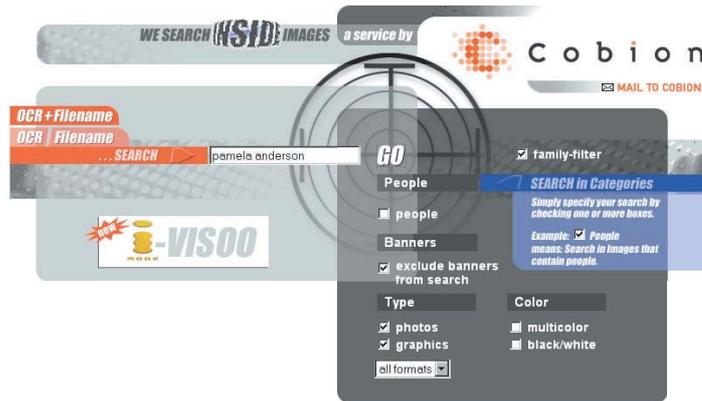


» бильному телефону на сервисные номера, где они анализируются. Результат произведенного анализа отсылается пользователю обратно по SMS.

Но это может быть намного проще: работая группа электронных медиатехнологий Фраунхофского института (AEMT) разработала Query by Humming. Вместо того чтобы мучить продавца дисков своими познаниями в музыке, в будущем достаточно будет только пропеть необходимую песню в микрофон, и магазин сразу сообщит название.

Суть техники в следующем: Query by Humming анализирует волны из отрывка песни и сравнивает результат с имеющимся аналогом в базе данных, в котором содержится дополнительная информация о песне: название, исполнитель, направление в музыке, вид такта или темп. Вся эта информация может распознаваться из аудиоданных автоматически.

Конкретные продукты с Query by Humming пользователи Сети могут ожидать в конце года: AEMT совместно с Deutsche Telekom разрабатывают модель сервера для сетевого пользования и для мобильной телефонии. Руководитель AEMT профессор Карлхайнц Бранденбург мечтает о том, чтобы пользователю нужно было только сказать название песни, которую он хотел бы слышать, а стереосистема должна была бы автоматически проиграть нужную мелодию.



◀ Сайт поисковой машины Cobion

Крепкий орешек: анализ движущихся картинок

Большие проблемы ожидают исследователей аудиоданных при анализе картинок. До сегодняшнего дня обычные системы описания картинок должны были быть «натренированы» на конкретный объект для поиска, в противном случае ничего не выходило.

Особенно гибкой оказалась Content Analysis Library компании Cobion. Удачность этой поисковой машины базируется на умных дешифровальных программах, которые измеряют и анализируют объекты, отталкиваясь от таких параметров, как размеры, расстояние, на которое они отстоят, и отношение друг к другу. К тому же Analyse-Tool различает тип шрифта, размер, цвет, позицию. Достаточно хорошо функционирует категоризация черно-белых и

цветных фотографий, графических изображений или рисунков.

Сама по себе удаётся выборка похожих картинок порнографического содержания. Система Cobion точно отличает кожу от текстиля цвета кожи и может даже определить местонахождение определенных пигментов. Так, например, по запросу «Памела Андерсон» фильтр выведет только те изображения, на которых у этой дамы в интимных местах есть одежда, пусть даже скудная. Cobion планирует распространить этот метод на анимированные графические форматы. Если такой инструмент станет ходовым товаром для видеопотока и медиафайлов, то ничего не будет стоять на пути всеобщей службы с информацией из всезнающей семантической сети — даже различные желания и привычки потребителя. **CHIP**

Круглосуточный сервис — всегда онлайн, больше не нужен поиск благодаря персонализированным услугам



2002

Сегодня — Поток информации в Сети угрожает умственным сферам деятельности

2003

UMTS — Скорость скачивания в мобильном Интернете до 2 Мбит/с будет возможна с использованием этой сетевой технологии

2005

IPv6 — Новый интернет-протокол даст уникальные IP-адреса всем подключенным к Сети устройствам

Будущее

Тим Бернерс-Ли: «У меня есть мечта. В первой части мечты Сеть становится одним из самых существенных инструментов для совместной работы людей. Во второй ее части и компьютеры начинают совместную работу. Они будут в состоянии анализировать все данные, находящиеся в Сети. Ежедневные приемы работы в производстве, бюрократии и т. п. будут взяты на себя компьютерами, которые будут связаны с другими компьютерами».