

НАВИГАТОРЫ INTERNET: КАТАЛОГИ → ПОИСКОВИКИ → ПОРТАЛЫ

Дмитрий Ландэ,

dwl@visti.net

Информационный центр «ЭЛВИСТИ»

В Internet насчитывается свыше 6,5 млн. Web-серверов (данные компании Inktomi по состоянию на июль 2000 года) и свыше двух миллиардов страниц, из которых 85 % размещены в США (отчет компании Cyveillance). По заявлению этой компании (<http://www.cyveillance.com>), если темпы роста (7 млн. новых страниц в день) сохранятся, то к 2001 году количество Web-страниц удвоится. По прогнозам, "центр роста" Internet сейчас уходит из США. Темпы роста объемов информационных ресурсов в Сети настолько большие (для сравнения, можно отметить - количество Web-серверов в 1998 году было еще 1 млн.), что задача обеспечения навигации в ресурсах Internet кажется далеко не очевидной. Сегодня каждый пользователь на собственном опыте чувствует в действии один из самых больших парадоксов Internet – «полезной информации в WWW становится все больше, но найти что-то необходимое все сложнее». Традиционные универсальные средства «учета» информационных ресурсов Сети – каталоги и информационно-поисковые системы уже сегодня не справляются с задачей. Эффективными оказываются лишь узкотематические (или региональные) каталоги и поисковики. Как гласит опубликованное в журнале Nature исследование, максимальный охват адресов в Интернете - 16% - сегодня обеспечивает поисковая система Northern Light, за ней - 15,5% - следует AltaVista, а на третьем месте - 11,3% - Hotbot. Таким образом, самые крупные информационно-поисковые системы в мире охватывают в своих индексах не более 20 % существующих ресурсов. Существуют удачные попытки построения интегрированных «метапоисковых систем», например, система Alltheweb (<http://www.usss.alltheweb.com/>), по прогнозам ее авторов, должна охватывать 100 % информационных ресурсов Internet (на сегодня она заиндексировала с помощью механизма Lycos уже 1 млрд. страниц), однако не опубликованы принципы его реализации и качественные характеристики поиска (полнота, релевантность).

Для достижения большого охвата информационных ресурсов средствами информационно-поисковых систем необходимо учитывать архитектуру всего Web-пространства, а именно этой информацией до последнего времени никто не владел. Близкой к реальности математической модели не существовало до 2000 года. Лишь в этом году специалисты компаний AltaVista, IBM и Compaq совершили прорыв, математически описав «карту» ресурсов и гиперсвязей существующего пространства World Wide Web. Исследования опровергли расхожее мнение, будто Интернет – это единое густое пространство. Проследив с помощью поискового механизма AltaVista свыше 600 млн. Web-страниц и 1,5 млрд. ссылок, размещенных на этих страницах, ученые пришли к выводу о том, что топология Web-пространства соответствует модели галстука-бабочки" (Bow Tie), состоящей из следующих компонент:

- центральное ядро, или узел галстука, составляют Web-страницы, взаимосвязанные так тесно, что, следуя гиперссылкам, из любой из них в конечном счете можно попасть на любую другую.
- 24% Web-страниц - это "отправные Web-страницы". Они содержат гиперссылки, которые в конечном счете ведут к ядру, но из ядра к ним попасть нельзя.
- столько же - 24% - "оконечных Web-страниц", к которым можно прийти по ссылкам из ядра, но нельзя вернуться назад.
- 22% Web-страниц полностью изолированы от центрального ядра: это либо "мысы", связанные гиперссылками со страницами любой другой категории, либо "перешейки", соединяющие две Web-страницы, не входящие в ядро, либо "острова", которые вообще не пересекаются с остальными ресурсами Internet. Единственный способ обнаружить ресурсы этой группы - знать адрес. Никакие поисковые машины не смогут найти эти острова, если они в прошлом каким-то образом не соединялись с другими частями Internet.

Исследователи обнаружили, что пропорции этих четырех категорий в течение многих месяцев оставались неизменными, несмотря на значительное увеличение общего объема Web-ресурсов.

Топология "Bow Tie" поясняет динамический характер Сети и позволяет получить представление о некоторых особенностях сложной организации WWW.

Традиционные средства навигации в Web-пространстве – это каталоги и поисковые системы. Первыми появились Web-каталоги, как психологически наиболее приближенные к западному мышлению. Каталоги зачастую не требуют от пользователя ввода какой-либо информации с клавиатуры – достаточно воспользоваться гиперссылками, чтобы найти необходимую информацию. Трудно представить традиционный «бумажный» каталог, содержащий несколько миллионов ссылок (не говоря уже о миллиарде). Точно так же трудно ориентироваться в электронном Web-каталоге, не используя дополнительных возможностей, главной среди которых является возможность ввода «своего» запроса с клавиатуры. Рост объема Web-ресурсов и усложнение синтаксиса языков запросов привел к появлению количества информационно-поисковых серверов в Сети. Сегодня наиболее развитые системы навигации в Internet обладают свойствами как Web-каталогов, так и информационно-поисковых серверов. Среди таких систем – уже мировые лидеры Yahoo (<http://www.yahoo.com>), AltaVista (<http://www.altavista.com/>), Northern Light (<http://www.nlsearch.com/>), Hotbot (<http://www.hotbot.com/>). В России лидирующее положение занимают системы Яндекс (<http://www.yandex.ru/>), Rambler (<http://www.rambler.ru/>), Aport (<http://www.aport.ru/>). В Украине первые Web-каталоги появились в 1995 году, а поисковики – в 1997. Сегодня известно свыше десятка украинских информационно-поисковых серверов и около пятидесяти Web-каталогов. Среди лидирующих систем можно назвать InfoReS (<http://el.visti.net/>), META (<http://meta-ukraine.com/>), NSearch (<http://www.search.kiev.ua/>).

Именно Web-каталоги и информационно-поисковые серверы (и их симбиоз) стали прародителями нового типа Web-сервиса – порталов – «ворот в Internet». Сегодня прослеживается эволюция порталов от поисковых машин и каталогов до самостоятельных, насыщенных информацией и самодостаточных сайтов. Порталы, как новые объекты WWW, возникли в 1998 году. Основная идея их создания заключалась в стремлении, наряду с возможностями навигации в Сети, предоставить пользователю максимальный уровень сервиса, сделать так, чтобы каждый сеанс своей работы в Internet он начинал именно с данного ресурса. При этом типичный портал пытается не только привлечь пользователей удобными средствами навигации в Internet, но и «удержать» их, предоставляя на своем ресурсе максимум необходимой и полезной информации. Портал представляет собой сайт, организованный как системное многоуровневое объединение разных ресурсов и сервисов. Как правило, такой сайт совмещает в себе разнообразные функции, предлагает разноплановые информационные ресурсы и различные услуги (поиск, рубрикаторы, финансовые индексы, информация о погоде и т.д.). С момента появления первых порталов, их основные функции существенных изменений не претерпели - это средства реализации навигации в Сети, общения, новостная часть, торговля и службы приложений. Портал можно рассматривать как сайт (или совокупность сайтов), обеспечивающий удовлетворение основных потребностей пользователей путем реализации услуг (сервисов) в следующих областях: информация, бизнес, общение, а также предоставления инструментария, необходимого пользователю для продвижения собственного контента в рамках портала. В соответствии с данным определением портал должен включать четыре основных типа сервисов:

- информационный сервис - все, что помогает найти (при необходимости) и получить информацию;
- сервис реализации бизнес-функций - все то, что ориентировано непосредственно на продажу товаров/услуг;
- инструментарий пользователя - все, что помогает ему создавать и продвигать свой контент в Сети, прежде всего, бесплатный хостинг и бесплатный e-mail, рейтинги, баннеры, «анонсировщики» и др.;
- сервис обеспечения общения (community) - все, что направлено на удовлетворение потребности в общении.

Различают «вертикальные» и «горизонтальные» порталы. Вертикальный портал - это видовой или тематический сайт, обладающий всеми качествами портала, ориентированный на один вид или тип информационного наполнения. Горизонтальный портал - это поливидовой и политематический сайт, обладающий всеми качествами Web-портала. Горизонтальный портал может включать в себя несколько видовых или тематических «вертикальных порталов».

При отсутствии до последнего времени четкого определения понятия «портал» и, в связи с инвестиционной привлекательностью данного направления деятельности в области Internet, порталами зачастую стали называть себя сайты, не удовлетворяющие некоторым, а зачастую и

многим признакам порталов. Конечно же, создание порталов требует долгосрочных инвестиций. При отсутствии в СНГ фондового рынка, на котором котировались бы акции Internet-компаний, инвесторам приходится ориентироваться на текущие финансовые показатели проекта и на прогнозы его развития в будущем. Несмотря на убыточность едва ли не любого Internet-проекта в первые несколько лет существования, некоторые из них котируются довольно высоко. По словам Евгения Голанда, президента Internet-службы PORT.RU (<http://www.port.ru/>), уже через несколько месяцев после запуска проекта он привлек инвестиции в размере 1 млн. долл.; при этом суммарная стоимость компании оценивалась инвесторами в 5 млн. долл. К концу прошлого года эта оценка выросла в пять раз.

Сколько же должно быть порталов в русскоязычной части Internet, сколько - в Украине? Многие аналитики считают, что для России, например, достаточно двух-трех порталов, больше не потребуется ни пользователям, ни инвесторам. Если не принимать во внимание качества российских «порталов», то уже сегодня эта цифра превышена в сотню раз – разработчики, называя свои сайты таким образом, ищут инвестиции. С другой стороны, мировой опыт показывает, что на 1000 сайтов должен приходиться один навигатор (каталог, поисковик, портал). На основании последних расчетов, на украинскую часть Сети должно приходиться 10-20 порталов. Что интересно, именно около 20 Web-сайтов в Украине позиционируют себя как «портал», зачастую «вертикальный», например «Украинский туристический портал» (<http://www.tour.com.ua/>), Торговый портал «Трейдер» (<http://www.trader.relc.com/>). Оба сайта весьма содержательны и заслуживают пристального внимания. Без сомнения, первый из них обладает большинством признаков вертикального (а точнее информационно-справочного) туристического портала. В отличие от этого проекта, большинство украинских «порталов» выполняют роль скорее «навигаторов» (зачастую каталогов) информационных ресурсов – моно- или политематических.

Вместе с тем, в Украине уже сегодня зарождается несколько серьезных проектов, один из которых развивает Информационный центр «ЭЛВИСТИ» (<http://www.visti.net/>), создавая мультипортал на основе восьми существующих собственных информационных сайтов, своих возможностей, как одного из крупнейших украинских ISP и возможностей информационных партнеров.