

NEW MEDIA – НОВАЯ ИНФОРМАЦИОННАЯ СРЕДА

*Александр Григорьев (gri@visti.net),
Дмитрий Ландэ (dwl@visti.net),
Информационный центр «ЭЛВИСТИ»*

I. INTERNET – СВАЛКА ИЛИ КЛОНДАЙК

Понятие New media уже традиционно принято сопоставлять с сетью Internet. Сегодня каждый пользователь New media 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. Было бы логичным, чтобы владельцы некоторых систем подобного типа, договорившись, пытались охватить лишь определенные «вертикальные» фрагменты Сети, совместно решая задачу полного охвата ресурсов и обеспечивая качественную навигацию в своих областях. Однако, такая модель утопична, а тенденции на рынке глобальных информационно-поисковых систем никак нельзя назвать радужными. Реалии таковы – новизна охватываемой информации падает, навигационные сервисы в основной своей массе не улучшаются, а количество самих глобальных информационно-поисковых систем (которые за редким исключением не стали порталами) стремительно растет.

Поэтому эффективность использования всех подобных систем достигается только в случае обращения к относительно стабильной части информационного пространства, но парадокс заключается как раз в том, что Internet в основном таковым не является. Более того, до последнего времени никто не владел информацией об архитектуре Web-пространства. А именно эти сведения крайне важны для достижения, с одной стороны, эффективного охвата информационных ресурсов средствами информационно-поисковых систем, а с другой стороны (со стороны пользователя), для обеспечения возможности получения действительно необходимой информации из сети. Близкой к реальности математической модели не существовало вплоть 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 цельное 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 адреса. Никакие поисковые машины не смогут найти эти острова, если они в прошлом каким-то образом не соединились с другими частями Internet.

Было обнаружено, 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. В последнее время в связи с этим исследованием появились идеи построения нетрадиционных средств "обратного" браузеринга в Сети, например реализующих переходы от Web-сайта на те Web-ресурсы, с которых имеются ссылки на него. Эта технология позволила бы находить новую информацию на основании «оконечных Web-страниц» модели Bow Tie. А именно такими страницами зачастую оказываются научные статьи, рефераты или отчеты о научно-технических работах.

С точки зрения обновляемости информации, все Internet-пространство можно условно разделить на две составляющие – стабильную и динамическую. Стабильная составляющая Сети содержит информацию «долговременного» плана, например монографии, галереи, коллекции или архивы. Динамическая составляющая включает постоянно обновляемые или новые ресурсы. Небольшая часть этой составляющей вливается затем в стабильную, в то время, как большая часть «исчезает» из Сети.

В свою очередь, информационные потребности пользователя можно условно разделить на две части – знание-понятийные и новостные. Очевидно, что первая часть потребностей в большей мере удовлетворяется стабильной составляющей Internet, в то время как новостные потребности могут найти свое удовлетворение только в динамической составляющей New media.

II. NEW MEDIA И СМИ

На сегодняшний день New media де-факто заняла место в ряду других средств массовой информации. Любое средство массовой информации, будь то печатное издание, радиостанция или телеканал, обладает своими техническими возможностями. Для разных видов СМИ - эти параметры разные.

Характеристиками потенциала печатного издания служат его тираж, формат, число страниц и т.д. Для радиостанции или телеканала это - частота вещания, мощность передатчика, область охвата. Сеть Internet обладает своим техническим потенциалом - пропускной способностью каналов, количеством подключенных компьютеров, их характеристиками и т.д. Реализация потенциала New media так же, как и в случае традиционных СМИ выражается в посещаемости, популярности, аудитории и ее направленности.

В качестве подтверждения важности роли Internet как средства массовой информации приведем пример – пожар на Останкинской телевизионной башне, который обусловил бурный всплеск интереса к Internet в России. То, что стало катастрофой для российского ТВ, стало заметным толчком в развитии Рунета. Число посещений новостной ленты РИА "РосБизнесКонсалтинг" (<http://www.rbc.ru>) 28 августа составило 6 млн., что в 6 раз превышает средний ежедневный трафик. Internet-газета "Lenta.ru" (<http://www.lenta.ru>) сообщила, что за тот же день было зафиксировано 63,5 тыс. посещений, что в 2 раза превышает обычный показатель. Однако аналитики заявляют, что сдвиг в сторону СМИ в Internet предшествовал пожару в Останкино. Они говорят о том, что спрос на сетевую информацию вырос в августе, подогретый интересом к трагедиям на Пушкинской площади и на подлодке "Курск". В результате, по мнению специалистов, рекламный рынок Рунета может рассчитывать с октября на прирост ежемесячного бюджета примерно на \$150 тыс., т.е. на 20-30%.

Сегодня принято различать два понятия - СМИ в Internet и сетевые СМИ. В свое время, одновременно с переносом в Internet зародился процесс создания в Сети изданий, электронные версии которых дополняли (а порой и заменяли) традиционные. СМИ в сети Internet зачастую представляют собой прямую репликацию традиционных средств массовой информации на Web-серверах. Лишь немногие средства массовой информации, даже имея свое "представительство" в Internet, смогли «перешагнуть» рамки традиционного представления своей информации и стать полноценными сетевыми СМИ. Сетевые СМИ представляют собой новый тип носителей информации, изначально ориентированный на Internet, учитывающий многие нюансы представления информации в new media. Как правило, выпуск традиционным СМИ полноценного сетевого варианта требует не только изменения форматов и формы подачи информации, но и определенной семантической корректировки материалов. Сетевым СМИ присуще два огромных преимущества – это оперативность и интерактивность. Вторая особенность подразумевает возможность самостоятельного «выстраивания маршрута» при чтении материалов издания, используя механизм гиперссылок или встроенные поисковые системы. Вместе с тем, бытовавший ранее миф о «миграции» читательской аудитории «бумажных» СМИ в new media и негативного влияния сетевых СМИ на популярность традиционных прототипов на практике не нашел своего подтверждения, более того, справедливо обратное.

Уже сегодня в Украине свыше 400 традиционных СМИ имеют собственные Web-серверы, и лишь десяток – действительно сетевое представление, обеспечивая для пользователей современный уровень сервиса. Как и во всем мире, на сегодня, СМИ представлены в украинском сегменте Internet, практически во всем своем многообразии – свое «виртуальное

представительство” в Сети имеют информационные агентства, газеты, журналы, радио и телевидение.

В настоящее время в Internet представлены все ведущие информационные агентства страны (ДИНАУ, УНИАН, ИНТЕРФАКС, УНИАР). Появились фото-серверы (ДИНАУ и УНИАН), появляются исключительно электронные версии информационных продуктов. Вместе с тем, еще имеется большой ряд проблем, стоящих перед отечественными агентствами – это низкая оперативность подачи информационных сообщений, отсутствие развитых средств работы с архивной информацией в Сети, отсутствие справочных и информационно-аналитических серверов.

В Internet представлено около 350 украинских газет и журналов. Из них 10% занимают архивы изданий, которые уже по разным причинам не выходят (“Всеукраинские ведомости”, “Независтимость” и др.). 10 % украинских изданий, представляют материалы в Internet до выхода в печать (публикации), 60% - регулярно обновляются. Часть изданий (15 %) – исключительно электронные. Большинство из представленных в Сети изданий предоставляют свободный доступ (полностью либо частично) к своим материалам.

Радио в украинском сегменте Internet представлено более чем 50 серверами, наиболее известные из которых «Гала-Радио» (Киев) и «Радио Люкс» (Львов). Хуже всех других видов СМИ в украинском Internet представлено телевидение – около 20 серверов. Кроме названного, в Internet представлены серверы издательств, электронные справочники и каталоги газет и журналов, программы радио и телевидения, сайты организаций, занимающихся проблемами СМИ. Один из самых полных списков украинских СМИ в Internet размещен по адресу <http://uamedia.visti.net/link/>.

III. Попытки интеграции ресурсов

Конечно, большинство СМИ, представленных в Internet, находят своего пользователя, однако, если рассматривать всю совокупность сетевых СМИ как некую общность по отношению к конкретному пользователю (или группе пользователей), то обнаруживается ряд проблем, связанных с полнотой, релевантностью и оперативностью получения новостной информации. Пользователь или группа пользователей зачастую часами “зависают” в Internet, обходя сотни сайтов с целью получения новостей по определенной тематике (или предметной области). В этом поиске традиционные каталоги и поисковые системы оказывают лишь косвенную помощь – они указывают адреса сайтов соответствующей тематики. Однако, ни одна из универсальных поисковых систем типа Lycos или Rambler не поможет в поиске актуальной новостной информации – период индексации таких систем составляет от недели до нескольких месяцев. Тем не менее, количество новостной информации в российском и украинском сегменте Internet превышает 10 тысяч записей в сутки. В России работы по интеграции новостей СМИ ведутся в Национальной электронной библиотеке (<http://www.nns.ru>), которая охватывает свыше 400 источников. Интеграция сетевых новостей на неплохом уровне выполняются в России на серверах “ЗАГОЛОВКИ.РУ” (<http://www.zagolovki.ru>) и “Аккумулятор новостей” (<http://news.battery.ru>), а также в Украине на сервере Аппарата Совета национальной безопасности и обороны

(<http://www.rainbow.gov.ua>) и на сайте проекта “Паук новостей” (<http://www.topnews.com.ua>).

Николас Негропонте (Nicholas Negroponte) из MIT еще несколько лет назад на страницах "Wired" (<http://www.wired.com/wired/>), что будущее принадлежит электронным агентам по сбору информации. Некоторые инструменты фильтрации информационного потока сегодня можно видеть в Internet на серверах названных выше проектов.

Вместе с тем, фундаментальные разработки в этом направлении до последнего времени не анонсировались. Обычно поиск, фильтрация и сбор информации в сети Internet сопряжены с необходимостью, во-первых, отвлечения соответствующих человеческих ресурсов и временных затрат, во-вторых, требуют достаточной квалификации персонала и, к сожалению, не могут учитывать всех особенностей структуры Сети и представления информации в ней. Это, в свою очередь, не делает полученную выборку информации из сети Internet репрезентативной.

При этом информационный поток, “потребляемый”, например, организацией из сети Internet носит, как правило, выраженную предметную окраску, характеризуемую областью интересов данной организации. Одним из вариантов сокращения общих расходов на сбор и фильтрацию информации из сети Internet, является выделение специального персонала для выполнения функций сбора, селекции и доведения информации. Однако поиск и предварительная обработка информации из сети Internet в ручном режиме представляет собой достаточно трудоёмкий процесс и не даёт желаемого эффекта.

Решение перечисленных задач возможно путём создания автоматических и автоматизированных систем сбора, фильтрации и анализа информации, - т.е. своеобразных интеллектуальных посредников между пользователем или корпоративной информационной системой организации и глобальной сетью Internet. Подобная система выполняет всю “черновую” работу по сбору и селекции информации из сети Internet и создаёт документальную базу данных, специфицированную предметной областью заказчика. Погружение информации в базу данных сопровождается классификацией и частичной структуризацией первичной информации. Для последующей информационно-аналитической работы конечному пользователю корпоративной информационной системы предоставляются эффективные средства навигации и поиска информации в созданной документальной базе данных.

Технология подобного типа реализована в Информационном центра «ЭЛВИСТИ», и базируется на использовании системы, в состав которой входят такие основные функциональные блоки:

- подсистема сбора информации;
- подсистема фильтрации и анализа информации;
- подсистема хранения и доступа к информации.

Подсистема сбора информации предназначена для автоматического обхода (сканирования) доступных информационных ресурсов сети Internet, которые представляют собой сетевые средства массовой информации Украины (в первую очередь) и других стран мира. В состав подсистемы сбора информации входят:

- Internet-робот, обеспечивающий автоматический обход серверов сети Internet в соответствии с инструкциями, заданными специфицированным URL-навигатором;
- модули преобразования (конвертирования) файлов в унифицированный текстовый формат;
- подсистема конфигурирования URL-навигатора.

Основные функции второй подсистемы – фильтрация, отбор и классификация информации. В состав подсистемы входят:

- модуль фильтрации информации (система избирательного распространения информации);
- функциональный классификатор-фильтр;
- модуль автоматической классификации информации.

Основой подсистемы хранения и доступа к информации является информационно-поисковая система InfoReS-XL. Она предназначена для накопления, хранения и сопровождения информационных материалов, полученных из сети Internet и уже приведенных к унифицированной форме, содержащей дополнительные классифицированные поля. Информационно-поисковая система также обеспечивает доступ к отдельным новостным документам по различным критериям и категориям, просмотр, выборку, поиск и отображение их содержимого в удобной форме. Предусматривается возможность обращения непосредственно из документа, погруженного в базу данных, к первоисточнику информации в сети Internet (при условии сохранения информации по соответствующему адресу в Сети).

Рассмотренная технология автоматизированного сбора новостной информации в сети Internet нашла свое применение при формировании системы электронных телеконференций ElVisti.Info (40 тематических разделов, получаемых абонентами 20 украинских ISP), а также получила внедрена в ряде организаций. В настоящее время в «ЭЛВИСТИ» ведутся работы по созданию на основе технологии Stream системы избирательного распространения новостной информации по сохраняемым запросам отдельных пользователей сети Internet.

IV. Порталы – за и против

Реальное состояние по охвату Internet и уровень навигационного сервиса большинства глобальных информационно-поисковых систем вынудили их владельцев идти на шаги, направленные на оказание дополнительных услуг, удовлетворяя зачастую не только «навигационные» потребности пользователей. Именно по этой причине в 1998 году возник новый тип Web-ресурсов – порталы – «ворота в Internet». Основная идея их создания заключалась в стремлении, наряду с возможностями поиска в Сети (далеко не всегда высокого качества), предоставить пользователю максимальный уровень сервиса, сделать

так, чтобы каждый сеанс своей работы в Internet пользователь начинал именно с данного ресурса. При этом типичный портал пытается не только привлечь пользователей средствами навигации в Internet, но и “удержать” их, предоставляя на своем ресурсе максимум необходимой и полезной информации. Как правило, портал 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. Обычно порталы охватывают четыре основных типа сервисов:

- информационный сервис - то, что помогает найти (при необходимости) и получить информацию;
- сервис реализации бизнес-функций - то, что ориентировано непосредственно на продажу товаров/услуг;
- инструментарий пользователя - то, что помогает пользователю создавать и продвигать свой контент в Сети, прежде всего, бесплатный хостинг и бесплатный e-mail, рейтинги, баннеры, «анонсировщики» и др.;
- сервис обеспечения общения - то, 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 («ворот в Internet”), на которых зачастую отсутствовала даже навигационная компонента. Сегодня в терминологии и сознании тех, кто строит и инвестирует порталы, закрепилось мнение, что их создание требует долгосрочных инвестиций. При этом инвесторам приходится ориентироваться скорее не на текущие финансовые показатели таких проектов, а на прогнозы их развития в будущем. Вместе с тем, несмотря на планируемую убыточность едва ли не любого Internet-проекта в первые несколько лет существования (напомним, первые порталы появились 2-3 года назад), некоторые из них котируются довольно высоко. Например, 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.

Наряду со сказанным, можно заметить что существует несколько популярнейших в мире информационно-поисковых служб, которые не позиционируют себя в качестве порталов, акцентируя внимание на свои ноу-хау

в области охвата Сети и качества поиска. Это прежде всего проект компании Nothern Light Technology (<http://www.nlsearch.com/>), разработчики которого предпочитает называть свой продукт механизмом «исследования» Web. В компании убеждены, что валить все в одну кучу не имеет смысла. Для того чтобы отыскать следы редкой книги, и получить данные о тенденциях в малом бизнесе, по словам разработчиков, нужны две принципиально отличающиеся поисковые стратегии. Nothern Light стремится улучшить качество поиска за счет индексации 5413 источников информации, предоставляемой информационными агентствами, коммерческими журналами и деловыми периодическими изданиями.

Пересматривает свои позиции и AltaVista, которая сокращает свой персонал на 25%, увольняя служащих, работа большинства из которых связана с медиа-порталом AltaVista. Сокращение персонала является следствием пересмотра компанией своей стратегии и концентрацией ресурсов в области создания поисковых Web-машин. Руководство AltaVista рассчитывает, что намеченные им действия позволят компании получить первую прибыль в январе 2001 г. Катализатором изменений послужило уменьшение объемов рекламы со стороны Интернет-компаний. AltaVista планирует и далее предоставлять услуги бесплатного доступа в Интернет в США и Европе, но основным направлением, останется поисковая служба.

Приведенные примеры относятся к компаниям, безусловно, вносящим значительный вклад в развитие Internet, поэтому они в большой мере отражают тенденцию специализации, присущую Сети. Вместе с тем, потребности пользователя не ограничиваются исключительно поиском информации в Web-пространстве. Если это учесть, то становится понятной сегодняшняя популярность порталов и инвестиционная привлекательность порталостроительства.