

ОТЛИЧИЯ НОВОСТНЫХ ВЕБ-САЙТОВ И НОВОСТНЫХ АГРЕГАТОРОВ

Бабин Егор Александрович

бакалавр, РУТ(МИИТ), РФ, Москва

Информационные ресурсы в среде Интернет, предоставляющие пользователю доступ к новостной информации можно разделить на две категории, основываясь на принципе их работы:

- новостные веб-сайты с ручным отбором новостей;
- новостные веб-сайты с автоматическим отбором новостей;

Первые отличаются тем, что у них в штате есть набор сотрудников, вручную отслеживающие новостные поводы, обрабатывающие их, ищущие источники, подтверждающие истинность информации, а также занимающиеся тем, что пишут для новостей оригинальные заголовки, описания, а также текст новости, полностью оригинальный для каждой статьи и конкретного ресурса.

В данной предметной области важное место занимает технология агрегации информации, позволяющая разгрузить оператора от рутинной функции первичного просмотра больших потоков данных, дав ему возможность максимально сосредоточиться на содержательном анализе.

В российском сегменте Интернета существует масса ресурсов, предоставляющих подготовленную сотрудниками новостную информацию. В число таких входят как СМИ свободного формата, так и официальные информагентства, такие как «Интерфакс», «РИА Новости» и т.д.

При этом некоторые из них предоставляют как исключительно новостную информацию в виде коротких оформленных заметок с заголовком, описанием и текстом новости, так и полноценных статей, связанных как с аналитикой ситуаций, описанных в новостях, так и на отвлечённые теме.

Рассмотрим примеры таких подходов: например, главная страница сайта vs.ru, где предоставлена комбинированная из новостей и статей информация, а также главная страница сайта информагентства «РИА Новости», где присутствуют исключительно новостные заметки.

Новостные веб-сайты с автоматической обработкой новостей напротив, не зависят от сотрудников, а самостоятельно формируют ленту новостей. При этом важно понимать, что сами по себе эти сервисы не являются СМИ, не создают новости и не принимают участия в их редактировании, но обязаны проверять достоверность приведенной информации перед её цитированием согласно Федеральному закону "Об информации, информационных технологиях и о защите информации".

Такие веб-сайты называются новостными агрегаторами.

Новостной агрегатор – это веб-сайт или приложение, которое сортирует информацию и группы источников согласно заданным алгоритмам и критериям отбора.

В число этих критериев входят: количество ссылок с других ресурсов на статью

(цитируемость), скорость появления информации (свежесть) и наполненность новости фактической информацией (информативность). Как правило, новостные агрегаторы работают в автоматическом режиме, то есть отбирают, оценивают и публикуют новость согласно заданным алгоритмам отбора и отображения, создавая таким образом так называемую новостную ленту, но также существуют ручные новостные агрегаторы, работой с информацией в которых занимаются люди.

Стоит отметить, что актуальность зачастую является превалирующим качеством в оценке значимости новости, в некоторых случаях являясь даже более высокой, чем достоверность. Например, в случае эскалации напряжённого вооружённого конфликта между двумя противоборствующими (политическими и/или иными) силами в рамках конкретного места действий, при котором обе стороны имеют свою политику в отношении предоставления информации сторонним наблюдателям, следящим за ситуацией, но непосредственно в ней незадействованным. В таком случае сводки и данные, поступившие от самых нижестоящих участников или непосредственных неаккредитованных наблюдателей событий могут формировать совершенно другую картину, нежели информация из официальных источников. Отсюда очевиден вывод о важности технологий отбора, обработки и представления информации.

-способность отбросить (отфильтровать) данные, не представляющие. Не все источники информации являются достоверными и надежными, что, вкупе с их обилием, формирует информационный шум. Важно понимать, насколько проблема информационного шума актуальна для современного общества. В условиях доступности и массовости средств получения информации заинтересованный пользователь может просто не найти нужные ему сведения, или получить информацию, не соответствующую действительности. Важнейшими свойствами агрегатора информации, определяющими, в конечном счете, его функциональную эффективность, являются:

- способность не пропустить в потоке данных информацию, действительно представляющую ценность;
- способность отбросить информацию, не представляющую интереса для пользователя.

Для достижения таких свойств необходимо применять специальные методы борьбы с информационным шумом:

- эффективному поиску информации способствует ее структурирование (по дате, по востребованности, по ключевым особенностям - автору, источнику и т.д.);
- для борьбы с дезинформацией, пользователю необходимо предоставить список альтернативных источников данной информации (в случае, если пользователь выбрал средство, которое не предоставляет эти источники, ему придется искать их самостоятельно). Вполне допустимо считать информацию достоверной в условиях её переизбытка, если её истинность подтверждается несколькими независимыми источниками.

В современном Интернете новостные агрегаторы играют очень большую роль – концентрируя на себе внимание посредством размещения актуальных новостей, они являются важным звеном процесса обмена информацией и имеют широкий охват аудитории. Прибыль от реализации подобных проектов напрямую зависит от количеств посещений данных веб-ресурсов.

Ситуацией, когда новостные агрегаторы получают выручку, при этом, фактически, не создавая ничего своего, лишь генерируя ленту из статей других изданий, заинтересовались многие органы исполнительной власти, в том числе в Европейском Союзе и в России.

Подобная ситуация ставит вопрос об авторском праве используемой информации, ведь, несмотря на то, что обычные, неавтоматизированные интернет-издания тоже говорят о вещах, в которых не принимают непосредственного участия, они в любом случае создают свой, уникальный контент, описывающий каждую конкретную ситуацию иными словами. Новостные агрегаторы же, в свою очередь, лишь пользуются их контентом, при этом,

естественно, давая ссылку на первоисточник.

В Европейском Союзе рассматривался вопрос о введении платы за контент для новостных агрегаторов, но в итоге было решено обязать новостные агрегаторы заключать договора с информационными ресурсами, имеющими более жесткую политику в отношении использования материалов с сайта.

Также обсуждались вопросы об ответственности новостных агрегаторов за информацию, ими предоставляемую. Особо важным моментом в данной дискуссии является то, что новостные агрегаторы не влияют на информацию и не редактируют её, предоставляя лишь ссылки на другие ресурсы.

В России этот вопрос был поднят на государственном уровне, шла речь о том, чтобы приравнять новостные агрегаторы к СМИ, как некоторые новостные сайты. Позднее этот вопрос был решен созданием специального реестра новостных агрегаторов в государственном ведомстве «Роскомнадзор». Нет сомнений, что в сфере государственного (а возможно и международного) регулирования последуют и другие действия, что в конечном счете, может значительно сузить общую информационную базу сети Интернет. Это говорит о необходимости выработки и отстаивания консолидированной позиции контент-провайдеров и IT- сообщества в целом.

В заключении имеет смысл сказать о перспективах дальнейшего развития в области получения, обработки и предоставления информации на различных веб-сайтах, в том числе новостных агрегаторах. Это развитие связано с совершенствованием как средств вычислительной техники, так и методов алгоритмизации, в том числе технологии нейросетей, построенной на основе искусственного интеллекта, эффективность которого напрямую зависит от вычислительных мощностей, предоставляемых для работы алгоритма. С учетом достаточно высоких темпов прогресса в этой сфере качество работы по сортировке и предоставлению пользователям новостного контента под их специфические интересы будет и далее возрастать, в то время как востребованность ручного отбора информации останется лишь в узкоспециализированных сегментах, в которых необходимо человеческое вмешательство для более грамотной оценки и вынесения суждений.