

Журналистский текст: особенности построения в условиях конвергенции и конкуренции

В эпоху активного развития технологий выпускнику факультета журналистики, не желающему остаться на обочине профессии, необходимо умение эффективного и оперативного использования преимуществ различных форматов подачи информации и понимание особенностей языка средств массовой коммуникации в современных условиях. В область его профессиональной компетенции входит не только овладение общими принципами и навыками формирования материала, но и знание хотя бы наиболее известных и общедоступных инструментов внешней аналитики, облегчающих процесс создания заголовков, текстов для более эффективной выдачи их в результатах поиска; и решение задачи информирования, нахождения взаимопонимания и координации поведения аудитории, чтобы не затеряться в информационном поле, чтобы знать, какие именно темы сегодня интересуют «своих» посетителей – для распределения степени актуальности имеющихся материалов и знания того, на какие темы стоит обратить внимание. Та же статистика употребления слов Wordstat от Яндекса, например, предоставляет данные по наиболее частым вариациям употребления ключевых слов, формулировкам поисковых запросов, и дает при этом возможность конкретизации региона, для которого собирается статистика.

Кроме того, необходимо помнить о том, что сайт зачастую не сможет заполучить интересующее его количество пользователей исключительно напрямую, по прямым заходам – да, можно рассчитывать на свою постоянную аудиторию, но можно и нужно постоянно привлекать новых посетителей, часть из которых впоследствии может стать постоянной и активной аудиторией. Способствовать этому может, в частности, комплекс мер поисковой оптимизации (SEO – search engine optimization). Меры эти направлены на поднятие позиций сайта в результатах выдачи поисковых систем: чем выше позиция, тем больше заинтересованных посетителей.

В основном, все факторы, влияющие на положение сайта в выдаче поисковой системы, можно разбить на внешние и внутренние. И как раз работу с внутренними факторами по оптимизации ресурса может и должен выполнять современный журналист – приведение текста и разметки страниц в соответствие с выбранными запросами, улучшение

качества и количества текста на сайте, стилистическое оформление текста, улучшение структуры и навигации, использование внутренних ссылок.

Любая поисковая система при оценке релевантности (адекватности содержания, степени соответствия введенному запросу) учитывает следующие параметры сайта:

- плотность ключевых слов – стоит учитывать, что алгоритмы поисковиков сейчас позволяют производить семантический анализ текста, чтобы отсеять поисковый спам, когда ключевое слово встречается слишком часто.
- индекс цитирования (ИЦ) – показатель поисковой системы, вычисляемый на основе числа ссылок на данный ресурс с других ресурсов сети. В простейшей разновидности индекса цитирования учитывается только количество ссылок на ресурс. Тематический индекс цитирования (тИЦ) учитывает также тематику ссылающихся на ресурс сайтов, а взвешенный индекс цитирования – популярность ссылающихся сайтов (также в большинстве случаев вычисляемую на основе индекса цитирования). До появления оптимизаторов сайтов индекс цитирования реально отражал популярность соответствующего ресурса в интернете. Первой крупной поисковой системой, начавшей активно использовать индекс цитирования, стала Google (алгоритм PageRank или PR). В русскоязычном сегменте интернета наибольшей известностью пользуется ИЦ Яндекса (Тематический индекс цитирования или тИЦ). тИЦ – технология поисковой машины, заключающаяся в определении «авторитетности» интернет-ресурсов с учетом качественной характеристики – ссылок на них с других сайтов. тИЦ рассчитывается по специально разработанному алгоритму, в котором особое значение придается тематической близости ресурса и ссылающихся на него сайтов.

Большое значение при ранжировании результатов поиска имеют заголовки страниц (теги <title>). Для попадания на первую страницу результатов поиска по низкочастотному запросу необходимо хотя бы частичное совпадение заголовка и запроса – вхождения слов запроса в заголовок страницы. При этом title – это не только технический атрибут страницы для поисковиков, но и то имя страницы, которое отображается в шапке окна браузера. И поскольку это имя видят еще и пользователи, оно должно быть, во-первых, удобочитаемым, во-вторых, максимально насыщенным ключевыми словами.

Благодаря своей востребованности SEO-методика стала достаточно популярной. Появились понятия «белой», «серой» и «черной» оптимизации – в зависимости от формальной разрешенности методов. В поисковиках появились и свои методы борьбы с «черной» оптимизацией. В частности, по первоисточникам отсекаются ресурсы, содержащие

большое количество материалов, заимствованных из сети (более 60%). Также отсеиваются и сайты, содержащие большое количество ссылок на другие ресурсы (более 20%).

При этом современный журналист обязан помнить о том, что ужесточается и принцип построения текста публикации: большое внимание уделяется заголовку и следующему за ним тексту, расположению и частоте ключевых слов сюжета в тексте. Здесь на помощь журналистике приходит и математика – ведь от фактического построения текста публикации напрямую зависит формирование аннотации – текста не длиннее 300 символов, который сопровождает ссылку, предоставляемую поисковиком; именно по ней определяется релевантность материала и принимается решение о переходе по ссылке. В частности, один из алгоритмов контекстно-зависимого аннотирования Яндекса основан на вычислении нескольких признаков запроса и лексикографической сортировке предложений по совокупности всех признаков. Алгоритм построения аннотации работает следующим образом: рассматривается первое предложение после сортировки, если это предложение по числу символов может публиковаться в тексте аннотации, то оно помещается в нее. Если же нет, то оно обрезается таким образом, чтобы длина его не превосходила 150 символов, а после обрезки добавляется в текст аннотации. Так же добавляются все леммы предложения в так называемое множество использованных лемм. Далее производится ещё одна пересортировка: идёт оценка, какие леммы запроса не вошли во множество использованных лемм, и предложения переупорядочиваются по убыванию суммы слов запроса, не вошедших в первое предложение. После этой пересортировки продолжается добавление предложения (с обрезкой его по необходимости), пока очередное предложение становится в разрешенную максимальную длину текста аннотации. При этом предложение пропускается, если разность мощности множества лемм предложения и мощности множества использованных лемм меньше, чем четверть мощности множества лемм предложения. Последний прием позволяет избежать добавления предложений, содержащих только те слова запроса, которые уже вошли в предыдущие добавленные предложения. Предложения, попавшие в аннотацию, сортируются в соответствии с порядком, с которым они шли в исходном документе.

Вышеперечисленные требования к построению текста публикаций для новых медиа весьма желательны и по другой причине: в связи с возрастающим обилием информации, предоставляемой различными ресурсами, многие пользователи предпочитают не посещать каждый ресурс непосредственно – это занимает слишком много времени – а пользоваться программами веб-агрегации, позволяющими подписаться на интересующие их RSS-каналы (Google Reader – сервис, предоставляемый вместе с аккаунтом Google, Яндекс.Лента – сервис Яндекса и так далее). Обычно с помощью RSS дается краткое описание новой информации, появившейся на сайте, и ссылка на ее полную версию. Экспорт в формате RSS

может пригодиться и для автоматической публикации ссылок на внешних ресурсах, в частности, так работают ставшие весьма популярными в последнее время твиттер-аккаунты. Публикация ссылок лично редактором интернет-ресурса в данном случае бессмысленна по затратам сил и времени, поэтому настраивается (на интернет-сленге «прикручивается») автоматическая публикация в Twitter по обновлениям RSS-канала: в ленте появляется заголовок новости и ссылка на материал. Сейчас уже уверенно можно говорить, что этот сервис микроблогинга видоизменил современный интернет. Для новостного издания подобный способ распространения информации становится еще одним каналом подписки, публикующим ссылки. Особенно актуальным это стало после того, как поисковики стали индексировать публикации в микроблогах.

Кроме того, в связи с общей тенденцией к сокращению размеров публикаций в современных медиа иногда необходимо разделять цельный сюжет на условные части – так, чтобы каждая из них могла существовать как отдельный новостной сюжет, не теряя смыслового содержания. Иногда для этого необходимо менять изначальную структуру текста, разработанную журналистом, и компоновать сюжеты по смысловым составляющим. В итоге исходный сюжет становится несколькими, более конкретными по тематике новостями, которые индексируются роботом-поисковиком по отдельности. Это, в свою очередь, в качестве различных новостных поводов, способно привлечь на сайт больше посетителей. В качестве заголовка, способного привлечь внимание читателя, может быть выбрана одна из основных цитат сюжета, которая в какой-то мере характеризует его в целом. В таком случае, в подзаголовке и первых абзацах текста даётся максимум информации, выполняющей описательную функцию для всего сюжета в целом – по принципу перевёрнутой пирамиды, для лучшей индексации поисковиками. К тому же разовое событие может обсуждаться достаточно долго, образуя отдельные рубрики с подборкой материалов на эту тему. В таких случаях заголовки и подзаголовки должны быть наиболее информативными, чтобы показать, что изменилось с момента публикации предыдущего текста, и содержать как можно большее количество слов, ставших ключевыми для этой темы – именно на них, скорее всего, будет строить новостной запрос пользователь.

Журналисту современного медиа нельзя забывать и о материалах постоянного интереса: тематическая информация – познавательная, справочная – является гораздо более «долгоиграющей» и может заинтересовать пользователя и через некоторое время после её публикации и потому, при грамотной редакции, даже давние по дате публикации сюжеты дают небольшой, но стабильный заход посетителей. Заголовки для таких сюжетов конструируются по принципу максимального подобия с содержанием предполагаемого поискового запроса. Для его определения можно опять же воспользоваться внешними

инструментами (Яндекс.Wordstat, с учётом региона). Ключевые слова таких текстов для лучшей индексации можно выделить дополнительно. К тому же можно рассчитывать на публикацию ссылок на такие материалы на тематических форумах, либо размещать подобные ссылки самостоятельно под видом обычного пользователя (сфера, подобная партизанскому маркетингу). Еще один полезный прием: некоторым материалам не нужно масштабное продвижение – достаточно отправной точки, и добровольная публикация ссылок, ведущих на страницы этих сюжетов, может распространяться подобно эпидемии. По этому принципу работает вирусная реклама, он же применим и для журналистики.

Таким образом, как уже отмечалось ранее, на построение и редактирование текста для современных медиа помимо филологических влияют как минимум математические, технические и технологические факторы, связанные с особенностями работы поисковиков. К тому же сегодня принципы работы определяются не только техническими факторами, значительно усиливается необходимость наличия как можно более полных и актуальных знаний в таких смежных областях как социология, психология, маркетинг, веб-дизайн и прочие. А необходимым условием успешной деятельности компетентного специалиста является опыт работы в интернете в различных направлениях, высокий уровень владения компьютером и новейшими информационными технологиями, постоянный поиск новых программ и ресурсов для улучшения организации работы. И потому можно смело говорить о том, что современная журналистика выходит за привычные нам рамки специализации, вбирая в себя новые профессиональные знания.