



ТЕХНОЛОГИИ МОНИТОРИНГА МАССОВОЙ ИНФОРМАЦИИ: от бумажного – к цифровому, от ручного – к автоматическому

Появление в 2000 г. электронной библиотеки Public.Ru, агрегирующей публикации в СМИ, стало ответом инициативной группы разработчиков ресурса на бурное развитие российского медийного рынка. У проекта была цель – избавить пользователей от необходимости искать сообщения по той или иной теме в тысяче разных источников. Предоставляя доступ к массиву цифровой информации, технологические возможности библиотеки позволяют клиентам самостоятельно, легко и быстро отбирать искомое и проводить его анализ.

Исходя из практики, в мониторинге СМИ нуждаются пресс-службы и PR-отделы бизнес-структур и государственных учреждений, PR-компании, маркетинговые агентства, исследовательские институты, вузовские и публичные библиотеки. А как показывает история, первая технология мониторинга СМИ была изобретена 130 лет назад нашим соотечественником, предпринимателем Генри Ромейке.

Автор



**Олег Васильевич
ТКАЧУК**

кандидат
социологических
наук, доцент
кафедры «Связи
с общественностью
и массовые
коммуникации»
Института
иностраных
языков МАИ

Генри Ромейке родился 17 ноября 1855 г. в Клайпеде (в то время город назывался Мемель и входил в состав Восточной Пруссии). По всей видимости, Генри был выходцем из увядающего древнего белорусского рода, родоначальник которого – Гурко (Гурий) Олехнович Ромейко, – был наместником смоленским в Витебском воеводстве (1539).

В середине XIX века Европа испытывала значительные экономические трудности, которые заставляли людей искать работу порой очень далеко от родных мест.

Оказавшись в 1879 г. в Париже¹, Ромейке обратил внимание, что женщина, работавшая в газетном киоске, подала одному из покупателей пачку из десяти газет, а тот протянул ей десять франков и ушёл, не получив сдачи, – за такую сумму в то время можно было оформить годовую подписку на несколько еже-

дневных газет. Киоскёрша рассказала любопытному Ромейке, что покупатель – известный художник, у которого только что прошла выставка его картин. По его просьбе она просмотрела все парижские газеты и пометила в них материалы, где упоминался художник. Это событие навело Ромейке на мысль заняться бизнесом – продавать людям вырезки из газет.

В тот же год он открыл своё первое предприятие в Париже, и местные журналисты сразу прозвали его «газетным паразитом». В течение нескольких лет всю работу он делал сам, имея небольшую прибыль.

В 1881 г. Ромейке в Лондоне договорился с оптовым продавцом газет *Curtice* об использовании непроданных экземпля-

1. A Unique business: Scanning the papers // *Southland Times*, Issue 11866, October 14, 1891, P. 3.

ляров. Так было открыто первое в мире Бюро газетных вырезок, сразу же получившее широкую популярность. О. Уайльд и многие другие публичные люди того времени регулярно получали вырезки газетных публикаций, в которых упоминалось их имя.

Вырезки аккуратно наклеивались получателями в новый модный атрибут гостиной того времени — скрап-бук — альбом вырезок из газет, который свидетельствовал хозяину и его гостям о степени публичного интереса к человеку.

Мода на коллекционирование газетных вырезок захлестнула Австралию, Старый и Новый Свет. В доме каждого публичного человека, в офисе каждой фирмы на видном месте красовались шикарные альбомы вырезок. Вырезки о самом себе коллекционировал и Ромейке.

Один из крупных концернов заказывал все упоминания в прессе, которые только можно найти, об испытаниях бронированной стали. Другой клиент желал получать уведомления о тех, кто собирается ехать за границу, чтобы попытаться продать им одежду для путешественников. Бюро Ромейке собирало газетные вырезки материалов о перчатках, пуговицах, зонтах, ювелирных новинках, новых отелях и ресторанах, новых способах применения целлулоида, собаках, новых учебниках, сигаретах, мехах, маяках, танцах, подвязках, чае, фольклоре, Армии Спасения, спицах, оправках и втулках, детских садах, светлом пиве, хлопковом масле, новых тротуарах и всех видах изобретений. Писатели и издатели покупали упоминания в газетах о своих новых книгах; торговцы произведениями искусства — о новых сделках и упоминаниях их картин; скульпторы заказывали вырезки о намерениях возводить памятники, чтобы своевременно принять участие в конкурсе на такую работу. Одна предприимчивая фирма собирала с помощью Ромейке информацию о регистрации браков, чтобы своевременно рассылать рекламные проспекты о своих детских колясках. Многие правительственные организации заказывали газетные вырезки с материалами, касающимися их деятельности.

В 1886 г. Ромейке перебрался в Америку. К этому времени уже была налажена работа его подразделений в Берлине, Лондоне и Париже. Первый офис в США Ромейке открыл, по видимости, в 1889 г., а год спустя предприятие переехало в самый центр Нью-Йорка — на Бродвей. В 1891 г. в его американском Бюро работали около пятидесяти чело-



public.ru
медиапоиск и анализ

Олег ВЛАДЕС
Генеральный директор интернет-библиотеки СМИ Public.Ru

Какую деятельность осуществляет электронная библиотека СМИ Public.Ru?

Компания специализируется на предоставлении информационно-аналитических услуг, проведении медиа-исследований и является владельцем интернет-библиотеки русскоязычных СМИ. Сегодня в фондах Public.Ru собраны публикации более 37 тысяч изданий: газет, журналов, информационных агентств, интернет-сайтов, а также транскрипты программ ведущих теле- и радиоканалов.

Из каких источников происходит комплектование фондов?

С 2004 г. мы работаем на условиях прямых договоров с издательствами, редакциями и поставщиками информации. В нашей системе каждый материал сопровождается ссылкой (гиперссылкой) на источник информации и данными об издании, авторе, дате публикации и пр. Условием нашего клиентского договора также является соблюдение авторских прав при использовании материалов из фондов интернет-библиотеки.

На основе лицензионных соглашений в нашей библиотеке представлены все значимые общественно-политические, деловые и специализированные издания. Загружаются издания из всех регионов России: газеты областного (республиканского, краевого) уровня; общегородские газеты; местные вклады центральные изданий; местные общеэкономические и деловые издания; пресса, отражающая экономическую, социокультурную, конфессиональную или иную специфику региона.

В интересах профессиональных пользователей стараемся соблюсти баланс: с одной стороны, предоставить им как можно более широкий охват СМИ, с другой — свести к минимуму количество источников, дублирующих информацию других изданий, так называемых перепечаток.

Какова технология обработки новых поступлений в библиотеку?

Материалы поступают в различных форматах. Большинство автоматически конвертируются в единый текстовый формат хранения. Материалы, поступающие на бумажном носителе или pdf, обрабатываются методом распознавания текста (OCR), после чего автоматически конвертируются в формат хранения.

Кто Ваши клиенты?

Ресурс Public.Ru часто используют для анализа репутации и эффективности PR-служб, для подготовки тематических и отраслевых исследований. Клиенты библиотеки — пресс-службы и PR-отделы бизнес-структур и государственных учреждений, PR-компании, маркетинговые агентства, исследовательские институты, университеты, публичные библиотеки. Среди на-



век. Они читали газеты, делали из них вырезки и отправляли их тысячам клиентов. В тот период Бюро еженедельно рассылало около 30 тыс. вырезок на самые разные темы. Доходы Бюро Ромейке позволяли ему работать уже не только «под заказ», но и впрок. Он начинает собирать и хранить все упоминания в прессе о каждом публичном человеке, чтобы продать эти архивы любому заказчику в любой момент. Бюро Ромейке стало описывать и каталогизировать все ежедневные газеты.

Политики начинают заказывать Ромейке материалы газет о процессе разработки и принятия законов. Он информировал послов о том, что писали газеты о них и их странах; иностранных консулов — о том, кто и что сказал о них в их отсутствие. Бюро поставляло Эдисону вырезки из газет о нововведениях в электричестве. Другие учёные получали вырезки о туннелях и новостях в области воздухоплавания.

Только 25% заказчиков Ромейке — публичные люди. Остальные — корпоративные клиенты — радиовещание, авиация, шоу-бизнес, автопроизводители, поставщики коммунальных услуг и торговые ассоциации.

В 1894 г. в Бюро Ромейке работали уже более 100 сотрудников². Около 60 молодых женщин просматривали 1900 ежедневных газет и 5000 еженедельников. Перед каж-

ших клиентов также службы безопасности компаний и консалтинговые агентства.

Сегодня во всём мире документы и материалы СМИ являются важнейшим информационным источником, отражающим историю новейшего времени. Историки и исследователи всё активнее вводят в научный оборот массивы документов СМИ и подвергают их серьёзным исследованиям. Например, при создании целой серии книг «Экономическая летопись России» её создатели обращались к материалам прессы прошлых лет из фондов Public.Ru. Авторы исследовали двадцатилетний период становления современной экономической системы после разрушения плановой экономики социализма.

Огромный интерес к нашим фондам есть у лингвистов. Новые явления, события, развитие общества порождают и новые термины, слова и устойчивые выражения. По нашим систематизированным архивам прессы специалисты изучают современный язык и процесс его неологизации: возникновение новых словоформ, индивидуально-авторские слова и выражения, новые фразеологизмы и метафоры.

Какие сервисы предоставлены пользователям?

Главный сервис электронной библиотеки — сквозной поиск с использованием ключевых слов по обширному архиву. Однако в проекте регулярно разрабатываются, тестируются и внедряются новые информационные сервисы в области статистической обработки и систематизации результатов поиска. Сейчас у пользователей уже есть возможность формировать собственные персональные коллекции статей, изданий или авторов; получать автоматическое уведомление о появлении новых статей на e-mail; проводить статистические измерения упоминаемости темы в СМИ и анализировать результаты замеров на графиках и диаграммах.

Какова структура компании? Назовите, пожалуйста, основные подразделения и службы.

У нас выделены такие стратегические направления развития проекта, как работа по формированию архива (Служба формирования фондов, Служба сбора информации), новые программные разработки и техническая поддержка проекта (Техническая дирекция, Служба технической поддержки), коммерческая деятельность (Коммерческий блок), деятельность по связям с общественностью (PR-служба). Общее количество сотрудников, занятых в Public.Ru, больше 100 человек.

Нужно отметить, что Служба сбора информации работает круглосуточно и без выходных — ежедневно обрабатывается и загружается в базу около 30 тыс. документов. Мы самостоятельно сканируем около 400 центральных изданий, обеспечивая 100%-ное соответствие текстов вышедшим номерам изданий. А благодаря Службе технической поддержки наши фонды доступны в любое время суток.

Можно ли сегодня рассчитать стоимость одного документа в мониторинге?

Мы коммерческая компания и финансируем себя сами, предоставляя платный доступ к собственным фондам. Цена документа в нашей базе зависит от условий сотрудничества с издательством, сложности предобработки документа — приведения его к единому текстовому формату. Средняя цена статьи составляет 15 рублей, но есть и документы стоимостью более 500 рублей.

*Над материалом работали:
Валерия Королёва, Варвара Осипова*

2. Mr. Romeike knows Greer // The New York Times, January 27, 1894.

дой работницей висел машинописный список клиентов и «ключевых поисковых запросов», в соответствии с которым следовало искать газетные публикации. Основная часть списка составляла до 7000 имён и слов. Дважды в день женщина-менеджер звонила в колокольчик и созывала работниц, чтобы вслух повторять имена новых клиентов и новые поисковые запросы. В газетах девушки галочками помечали значимые публикации, которые затем вырезала острыми, как бритвы, ножами группа мальчиков. Ещё одна бригада молодых женщин сортировала вырезки по ячейкам с именами клиентов и другими поисковыми запросами.

Речь шла о колоссальном объёме работы. Например, Комиссия по празднованию двухсотлетия со дня рождения Джорджа Вашингтона только за один месяц получила от Бюро Ромейке 76206 вырезок, заплатив по 4 цента за каждую. При этом стандартная ставка Бюро составляла 7,5 центов для партий до 100 вырезок, собранных в месяц, 5 центов — за первые 2000. За хранение информации о клиенте «до востребования» Ромейке брал по 5 долларов ежемесячно. После двухлетнего испытательного срока работницы получали по тем временам значительную зарплату — около 20 долларов в неделю. Квалификация и опыт сотрудников не предохраняли от ошибок: довольно часто клиенты получали из Бюро вырезок материалы, не имевшие к ним никакого отношения.

Какое-то время термин *rotei- ked* настолько вошёл в моду, что означал более общее понятие — «быть собранным в альбомы».

Ромейке, несмотря на свой колоссальный коммерческий успех, почти не оставил наследства, — жил на широкую ногу. Он был членом самых популярных клубов Нью-Йорка того времени: «Лотос-клуба», «Пресс-клуба» и «Клуба 13».



Предприятие, основанное Ромейке в Лондоне, успешно работает до сих пор³, — сейчас оно называется *Cision*.

Бюро газетных вырезок в конце XIX — начале XX вв. открываются в крупных городах повсеместно в Америке, Европе, Австралии.

В 1901 г. и в Петербурге появилось «Общество трудовой помощи образованным лицам», более известное как первое в России Бюро газетных вырезок⁴.

Его работа заключалась в просмотре периодической печати и подборе газетных статей, на которые подписывались учреждения и частные лица. Любой желающий, внеся в задаток 1 рубль, получал почтой статьи по интересующим его вопросам. За каждую вырезку абонент платил от 5 до 13 копеек в зависимости от размера статьи.

В Бюро работало 14–15 человек, получавших ежедневно по 60–80 копеек жалования⁵. Вначале Бюро имело 15 постоянных клиентов, но уже в течение первого года работы их число возросло до 122. Бюро ежедневно получало

150–200 газет и журналов. Некоторые газеты выписывались, но большинство жертвовалось различными учреждениями и частными лицами. 47 редакций доставляли в Бюро свои издания безвозмездно; бесплатно получались 3 английские и 6 немецких газет.

К началу революционных событий 1917 г. и в российской сыскальной полиции была отработана технология систематизации газетных вырезок об интересующих полицию лицах и событиях⁶. В коллекции Толстовского музея РАН, например, хранится 20-томное собрание вырезок из газет по поводу ухода Л.Н. Толстого из Ясной Поляны (1910) и его смерти, сделанных Осведомительным бюро при Министерстве внутренних дел.

Таким образом, технологии «ручного» мониторинга СМИ в целом были отлажены уже в конце XIX в. и практически без изменений использовались в разных странах до появления в последнем десятилетии XX в. цифровых технологий хранения, обработки и передачи массовой информации. ■

3. <http://intl.cision.com/>

4. Бюро газетных вырезок // Известия по литературе, наукам и библиографии, 1901 г. Цит. по: Наука и жизнь, №1, 2001.

5. Там же.

6. Малыгин С.С., Чечётин А.Е. Основы оперативно-розыскной деятельности: курс лекций. Екатеринбург, 2001.