

Предисловие

Эта книга для тех, кто интересуется методами получения новых знаний на основе анализа современного информационного пространства, а также способами обработки информационных потоков с целью выявления тенденций, новых понятий, феноменов, взаимосвязей.

Одно из определений знаний, которое дает энциклопедический словарь Webster, следующее: состояние осведомленности о чем-то или обладание информацией. Именно эта трактовка знаний наиболее близка к проблематике данной работы. Объем данных, из которого приходится выискивать крупинки необходимой, актуальной, готовой к немедленному использованию информации для решения проблем, обуславливает актуальность и значимость самого процесса поиска знаний.

Если знания — это сила, то сегодня первоочередная задача — найти эту силу. При этом поиск знаний, в отличие от простого поиска информации, при котором зачастую не учитывается семантика запросов, должен предоставлять пользователю только действительно актуальную информацию, наиболее точно соответствующую его потребностям, и вместе с тем адекватную исходному запросу. Если при обычном информационном поиске пользователь в конечном итоге знает, что он может получить, то при поиске знаний он должен получить нечто до сих пор ему неизвестное и познать его.

О сложности такого процесса говорит, например, недавнее исследование, проведенное фирмой Reuters среди 1300 менеджеров, которое показало, что “менеджеры чувствуют, что не могут эффективно работать без получения большого объема информации, но эта тяжелая загрузка данными, часто не имеющими никакого отношения к делу, снижает эффективность их работы и препятствует нормальному функционированию корпоративной машины”. Это состояние было названо “синдромом информационной усталости”, что свидетельствует об избытке информации и недостатке знаний. Из опрошенных фирмой Reuters менеджеров, 38% утверждают, что “тратят много времени, пытаясь найти нужную информацию”. По оценкам экспертов, около 79% журналистов обращаются к Internet в поисках новостей и лишь 20% находят ту информацию, которая им необходима. Все они на самом деле ищут именно знания.

В последнее время о поиске знаний пишут достаточно много. Появилось новое направление в обработке текстовой информации — “глубинный анализ текстов” (Text Mining). Это направление, скорее технологическое, чем научное, включило в себя все реальные, реализуемые на практике результаты исследований в области контент-анализа и компьютерной лингвистики, которая, как и теория баз знаний, интенсивно развивалась в 70–80-е годы прошлого века.

Сегодня прагматичные подходы, свойственные технологии Text Mining, могут применяться как студентами при написании обзорных курсовых работ, так и маркетологами при анализе рынков, политиками, бизнесменами, учеными — всеми, кто активно участвует в современных информационных, политических и бизнес-процессах.

Методы Text Mining уже используются в таких основных областях, как:

- политические исследования — геополитика, анализ предвыборной и выборной ситуации, деятельность партий, общественных организаций, отдельных политических деятелей и т.д.;
- конкурентная разведка — обобщенный анализ деятельности конкурентов, их PR-активности, клиентской базы;
- анализ рынков — выявление основных тенденций в производстве и потреблении товаров и услуг определенных видов, в политике фирм, участвующих в рынках, ареалах;
- анализ новых технологий — в различных сферах науки, бизнеса, безопасности;
- образование, культура.

Несмотря на то что книга ориентирована на широкий круг читателей, интересующихся современными информационными технологиями, хочется верить, что она будет также полезна и аналитикам, которые с помощью методологии Text Mining или отдельных ее компонентов смогут повысить эффективность и качество своей работы.

Ждем ваших отзывов!

Вы, уважаемый читатель, и есть главный критик и комментатор этой книги. Мы ценим ваше мнение и хотим знать, что было сделано нами правильно, что можно было сделать лучше и что еще вы хотели бы увидеть изданным нами. Нам интересно услышать и любые другие замечания, которые вам хотелось бы высказать в наш адрес.

Мы ждем ваших комментариев и надеемся на них. Вы можете прислать нам бумажное или электронное письмо либо просто посетить наш Web-сервер и оставить свои замечания там. Одним словом, любым удобным для вас способом дайте нам знать, нравится или нет вам эта книга, а также выскажите свое мнение о том, как сделать наши книги более интересными для вас.

Посылая письмо или сообщение, не забудьте указать название книги и ее авторов, а также ваш обратный адрес. Мы внимательно ознакомимся с вашим мнением и обязательно учтем его при отборе и подготовке к изданию последующих книг. Наши координаты:

E-mail: info@dialektika.com
WWW: http://www.dialektika.com

Адреса для писем:

из России: 115419, Москва, а/я 783
из Украины: 03150, Киев, а/я 152