

ПРИМЕНЕНИЕ СВОБОДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ В АНАЛИТИЧЕСКОЙ РАБОТЕ

Гальченко Максим Иванович

Гущинский Александр Геннадьевич

Санкт-Петербургский государственный аграрный университет

Работа на должности «маркетолог-аналитик», «аналитик» предполагает систематическое использование баз данных, электронных таблиц, специализированных инструментов.

Естественно, что приходящий на рабочее место специалист, в силу полученных им навыков, будет стремиться использовать то ПО, которое он видел и применял ранее в процессе обучения. Однако, с точки зрения предприятия, маркетолог-аналитик, аналитик — это человек не участвующий напрямую в получении прибыли и комплектация его рабочего места соответствующим ПО не носит первоочередной характер. Если говорить об офисных пакетах, то тут траты вполне понятны и, скорее всего будут сделаны, но если говорить о специализированном аналитическом ПО — большие траты совершенно не очевидны, особенно на предприятиях малого и среднего бизнеса. Естественным выходом из создавшейся ситуации является применение свободного программного обеспечения.

Так как отдел аналитики, маркетинга работает, в большинстве случаев, на обеспечение внутренних потребностей в информации замена MS Office на свободные аналоги (OpenOffice.org, LibreOffice, Lotus Symphony) представляется вполне адекватным шагом. С учетом того, что в РФ 21.12.2010 принят ГОСТ Р ИСО/МЭК 26300-2010 «Информационная технология. Формат Open Document для офисных приложений (OpenDocument) v1.0», «фродной» формат свободных пакетов становится единственным стандартизированным. К тому же, данный формат достаточно полно воспринимается MS Office 2007 и старше.

Работа с базами данных на уровне запроса на выбор вполне осуществима с помощью компонента Base, входящего в состав OpenOffice.org и LibreOffice. В случае использования в качестве СУБД MySQL значительный объем работ с базами может быть вынесен в MySQL Workbench.

Исторически подготовка студентов к работе со статистическим программным обеспечением ведется в хорошо известных проприетарных пакетах, таких как SPSS и Statistica. Ежегодно компания Rexer Analytics проводит интернет-опрос пользователей ПО для интеллектуального анализа данных (data mining). Основные результаты опроса 2010 года были доложены в докладе «2010 Data Miner Survey Highlights » на конференции «Predictive Analytics World», презентация которого находится в открытом доступе. В опросе приняли участие 753 респондента, основная часть которых применяет методы интеллектуального анализа в CRM, маркетинге, финансах (60%), то есть — интересующей нас области. Опрос показал, что в среднем аналитики используют 4-5 инструментов в своей деятельности. В общем по выборке

наиболее популярным инструментом, применяем хотя бы от случая к случаю является язык статистического программирования R (43%), около четверти-пятой части (20%-25%) исследователей используют такие инструменты как WEKA, RapidMiner, KNIME. Из проприетарных пакетов традиционно выделяются SAS, IBM SPSS, Statistica, которыми пользуются около трети респондентов. Стоит отметить, что пользователи, в целом, удовлетворены применяемыми ими свободными пакетами, а в случае с R — 80% пользователей намерены продолжать использовать данный язык на протяжении последующих трех лет.

Из указанного выше набора data mining пакетов стоит отметить KNIME, который в первый раз вошел в список используемых инструментов в указанном ранее отчете, сразу встав на одну ступеньку с RapidMiner.