1нформационные технологии

Смена фокуса

В 2015 году Gartner убрал термин Big Data из своего графика "Нуре Cycle for Emerging Technologies", который показывает отношение бизнеса к самым обсуждаемым новым технологиям. Но это не значит, что класс решений для работы с большими данными перестал существовать. Они просто перестали быть предметом обсуждений — бизнес перешел к их повседневному использованию.

сегмент рынка —

Как показывает исследование Gartстировать в инструменты работы с большими данными стали больше, но реже. Фокус переместился непосредственно с больших данных на бизнес-проблемы, которые они могут решить. В 2016 году 48% компаний в мире инвестировали в эти инструменты, это на 5% больше, чем в 2015 году. Однако число компаний, которые планируют выделить бюджеты на это направление в ближайшие два года, сократилось с 31 до 25%.

По мнению Ника Хойдекера, директора по исследованиям Gartner, такие изменения свидетельствуют о том, что проблема теперь не в самих больших данных, а в том, как их используют. "Хотя организации поняли, что big data — это не какие-то специфические технологии, им теперь необходимо перестать думать о больших данных как о самостоятельной

Идея в том, что big data — это набор технологий и подходов к управлению данными для проведения анализа, применимого для решения самых разных задач. Директор по исследованиям Gartner также добавил, что успешное применение этих инструментов зависит от подхода, который использует компания. Целостный подход включает в

димых данных и инфраструктуры.

В России активные пользоватеner, проведенное в 2016 году, инве- ли big data — это в первую очередь финансовые компании. Президент и председатель правления Сбербанка Герман Греф на годовом общем собрании акционеров Сбербанка рассказал о том, как компания использует подобные инструменты: «Большие данные — это то, что позволило нам прогнозировать отток абонентов и сорадикально изменить процесс управления рисками. Банк использует порядка 600 моделей для оценки самых различных видов рисков: кредитных, труктурированных данных, а имениспользование больших данных из внешних источников в кредитном анализе. Это повлияло как на качество, так и на скорость принятия решений» На анализе больших данных построен, в частности, продукт для малого бизнеса «Смарт Кредит».

> правления ВІ компании «Аплана. Центр Разработки» (ГК АйТи) говорит, что big data также помогает банкам решать задачи кредитного скоринга, сегментации клиентов с целью формирования таргетированных предложений. "Например, мы построили и поддерживаем клиентское хранилище данных с помощью которого «Попродуктов и цифровых сервисов. Еще

персонала, использование необхо- ентам-физлицам. Зная уже достаточно много о клиентах и их предпочтениях — что и как они потребляют, можно точечно рассылать им рекламу", — рассказывает она

По словам Анны Михеевой, в ритейле с помощью big data можно прогнозировать спрос и, как следствие, оптимизировать складские запасы. В телекоммуникационном секторе здавать предложения новых продуктов и услуг. Также со стороны телекомов возрастает интерес к анализу несоперационных и т.д. Банк расширил но — к выявлению ценной информации в накапливаемых компанией до-

Константин Кутуков, директор по маркетингу группы «Астерос», уверен, что выиграть от использования этих технологий могут все без исключения отрасли, но морально и техно-Анна Михеева, Руководитель на- логически еще не все готовы к ее внедрению. "Основными потребителями big data являются в первую oчередь компании финансового сектора, телекома, торговли и, что особенно важно, государство" — говорит он.

По словам Константина Кутукова, коммерческие заказчики используют обработку больших объемов данных, например, для анализа платежеспочта России» использует информацию собности клиентов, потребительскоо своих клиентах для создания новых го поведения и рыночных трендов. Это помогает повышать конкурентоодно из направлений использования способность и быстро адаптировать себя сосредоточенность на бизнес- Від Data в «Почте России» — прямая предложения под конкретный сегрезультатах, вовлечение опытного адресная рассылка (Direct Mail) кли- мент заказчиков. Он приводит при-

мер: "У операторов большой тройки есть подразделения, специализирующиеся на работе с big data, причем они призваны решать в первую очередь задачу генерации дополнительной выручки. То же самое можно сказать о банковском секторе: решения по анализу больших данных внедрены в Сбербанке, Газпромбанке, ВТБ24, Альфа-Банке, ФК «Открытие» и мно- накоплены большие объемы трангих других. Ритейл (X5 Retail Group, «Глория Джинс», «Юлмарт», «Лента», «М.Видео» и др.) также находится в числе первопроходцев рынка big data начиная от анализа покупок, оптимизации логистики и номенклатуры закупок до раскладки товаров и управления программами лояльности".

В госсекторе, по заявлению Константина Кутукова, применение этих технологий может значительно повысить качество управления и оказания услуг населению, но пока используются мало. "В числе лидеров злесь ФНС. Аналитический центр правительства России, Пенсионный фонд, Правительство Москвы, Фонд обязательного медицинского страхования. Отдельным блоком стоят органы безопасности, для которых big data уже давно служит инструментом раннего выявления и предотвращения угроз, что особенно важно в нынешней ситуации, — говорит он. — Реализация концепции «умных городов» и развитие цифровой медицины также окажут большую поддержку внедрению технологий и повысят качество жизни и социальной инфраструктуры страны".

Если говорить о поставщиках решений для работы с big data, то, по словам Кутукова, на российском рынке преобладают пока западные ке квалифицированных кадров вендоры, чьи продукты основаны на отраслевых стандартах. В их числе Oracle, Microsoft, SAP, Teradata, EMC, IBM и другие. Но Кутуков уверен, что ласти больших данных требует пра-

рошие заделы и реальные шансы на успех в этом продуктовом сегменте.

Взять след

По мнению Анны Михеевой, процесс освоения этих технологий будет продолжаться по тем направлениям, которые пока недостаточно развиты. Например, компании, у которых уже закционных клиентских данных, пока еще не научились обрабатывать их в реальном времени. А это может пригодиться для предоставления клиентам персонализированных предложений в момент обращения, для выявления подозрительных операций. По словам госпожи Михеевой, повышенный интерес к таким решениям наблюдается со стороны заказчиков из финансовой сферы. Ритейл и производственная отрасль также демонстрируют рост интереса к инструментам для интеллектуального анализа имеющихся больших данных с целью совершенствования системы прогнозирования спроса, оптимизации объема производства и складских запасов с учетом множества факторов. "Для решения подобных задач мы используем платформы с открытым кодом Hadoop/Kafka/Spark, промышленные решения на основе СУБД PostgreSQL, продукты российского производителя Basegroup Labs, ETL платформу Pentaho Data Integration и многое другое. Мы ожидаем роста объема портфеля подобных заказов примерно на 40% в этом году", — добавляет она.

Константин Кутуков считает, что одна из ключевых проблем при освоении big data заключается в нехват-Data Scientists. "Специалистов такого профиля мало на рынке, а качественная реализация любого проекта в обу российских игроков тоже есть хо- вильной постановки задачи, грамот-

ного управления и специфических знаний" — объясняет он. В числе других препятствий он отмечает необходимость создания качественно новой среды для хранения, обработки и передачи данных, которая должна обеспечить возможность быстрой потоковой обработки (in-memory) разнородной информации. То есть помимо аналитического программного обеспечения необходим скачок в развитии инфраструктуры. "Отчасти это реализуется в рамках концепции «озер данных», когда централизованно агрегируются и обрабатываются большие объемы информации из множества источников" — поясняет

Необходимо решить и вопрос безопасности данных и соблюдения конфиденциальности — для работы с big data необходимо предусмотреть комплексное обеспечение защиты и обеспечить высокий уровень надежности систем. По словам Кутукова, препятствием для широкого освоения технологий big data в России остается и отсутствие у большинства компаний практики накопления больших объемов данных и их низкое качество.

"Тем не менее, в России уже появились операторы больших данных, аккумулирующие и обрабатывающие их в соответствии с потребностями заказчика, — говорит Константин Кутуков. — Это позволяет компаниям из разных отраслей быстрее и дешевле интегрировать инструменты big data в свою деятельность. Кроме того, технологии обработки больших данных являются фундаментом для развития систем машинного обучения и составным звеном технологий ИИ".

Применение алгоритмов глубокого обучения (maching learning) и искусственного интеллекта — следующий этап развития технологий

Светлана Рагимова

«Технологическая революция в сфере работы с данными открыла новые рынки»

— теория и практика —

Операторы связи, как и банки, первыми осваивают новые технологии. Директор департамента стратегического маркетинга МТС Леонид Ткаченко на конференции Teradata Universe 2017, которая прошла в Ницце в апреле. рассказал редактору «Ъ-Информационные технологии» Светлане Рагимовой о том, как анализ больших данных помогает повышать эффективность и увеличивать выручку компании.

— Какой опыт у МТС по работе с большими данными?

— Мы работали с большими данными всегда: у нас миллионы абонентов, и один из основных бизнес-процессов любой телекоммуникацион- туации, когда в салоне три человека, прибыльным. Сейчас мы поставиной компании — это биллинг. Даже просто посчитать, кто сколько платит компании, — это уже по факту работа с большими данными.

Но за последние три-пять лет произошла технологическая революция в сфере работы с данными, которая породила инструменты, позволяющие с ними работать по-новому, и, по сути, открыла для нас новые рынки.

Мы стали смотреть на эту тему иначе: не только базовые вещи делать, а создавать продукты, используя данные, которыми обладаем в силу характера нашего бизнеса.

— Приведите примеры, как это на практике может выглядеть? Знаю, что пробовали инструменты социальной аналитики, что еще?

— Да, мы используем до сих пор Social Analysis, чтобы анализировать социальные связи абонентов и выявлять лидеров мнений. Через них можно более успешно продвигать определенные продукты.

Кроме того, у нас есть очень много разных идей по монетизации данных. Приведу пример. Есть продукт по оптимизации рабочих графиков наших салонов. У нас порядка 5 тыс. магазинов, в которых работает огромное количество людей. При этом расписание рабочего времени до сих пор делалось, по сути, вручную, начальниками салонов. Сейчас технологии развились, мы можем собрать данные по всем трансакциям клиентов в одном месте: все визиты в магазин, покупки, обращения. В итоге мы видим, сколько реально нужно сотрудников в каждом салоне в каждый момент в любой день недели. Не нужно, конечно, составлять уникальное расписание под каждый день, достаточно понимать циклы, сколько в конкретный месяц нужно людей по рабочим часам. И вот сейчас мы меняем бизнес-процессы, чтобы оптимально распределить нагрузку в магазинах. В результате, с одной стороны, мы избавляемся от очередей, с другой — от си- пласт аналитических метрик, кото- лиза потоков передвижения потен-



а работает только один, а значит, экономим на фонде оплаты труда.

инструментов Big data — сегментация клиентов для таргетинга рекламы. Как имеюшаяся инфор- наберем огромный штат, а потом по**мация о потребителях, которой** смотрим, окупится это или нет». Мы владеют операторы, может использоваться в этой сфере?

данные, которых нет у интернет- ние оказывается успешным, инвескомпаний. Это в первую очередь информация об офлайн-активностях. рекламных сообщений. Например, сейчас все, кто пытается сегментиро- 500 млн руб. вать клиентские базы для каких-ливека. Допустим, человек посмотрел лиарда? преследовать, хотя он может быть давно это уже купил. Или знают, что зывают вам рекламу кормов для собак. Так сейчас делают все, так устропания не научилась выявлять наиболее релевантный момент для показа этой таргетированной рекламы. Ведь интересы человека непостоянны, они меняются. Важно не просто департаменту инновационных техпонять, что вам нужно. Правильным подходом будет определить, когда это нужно. Здесь операторы могут быть крайне полезны, поэтому мы сейчас работаем в том числе в этом направлении.

— **Каким образом вы пытаетесь** ствует уже несколько лет, а значит, решить эту задачу? Вводите до- его эффективность очевидна для заполнительные признаки? Вклю-

чаете в анализ геотеги? вые признаки к традиционным, принятым на рынке. У нас два слоя признаков для сегментации. Базовые — те, которые мы просто снима- ты мы реализуем и для коммерчеем с сети, как всю жизнь это делали: ских компаний, в том числе помокуда абонент ходит, какие звонки совершает, какие сайты посещает. Есть стить точки продаж, исходя из ана-

рые мы рассчитываем на основе базовых. К примеру, мы понимаем, что, если два человека много друг другу звонят и ночуют в зоне действия одной базовой станции, это одно домашнее хозяйство. Если прекращают звонить друг другу в пятницу вечером — признак, что вместе живут с большой вероятностью. У нас много таких аналитических метрик, построенных на базовых данных.

— Как у вас выстроена работа в департаменте? У ваших коллег, например, в отделе анализа больших данных может быть до сотни человек.

— Изначально у нас в команде работали 35 человек. Уже по результатам первого года нам удалось не просто окупить инвестиции в анализ данных, но и сделать это направление ли еще более амбициозные цели по выручке подразделения и расшири-— Одна из областей применения ли команду до 50 человек. Мы не работаем по принципу «сейчас купим самое дорогое решение у вендора и идем небольшими шажками, пробуя новые технологии, применяя их — У операторов есть уникальные на конкретных кейсах. Если решетируем в развитие проекта — это касается и финансовых инвестиций, Кроме того, мы сейчас пытаемся пе- и кадровых ресурсов. В результаресмотреть сам подход к доставке те по итогам 2016 года совокупный эффект от Big data в MTC превысил

— На каких именно проектах бо целей, исходят из интересов чело- удалось заработать эти полмил-

в интернете что-то про Грецию, и ре- — Порядка 80% принесли внутренклама отелей, туров начинает его иние продукты, которые повышают эффективность работы группы МТС, и только 20% — это то, что мы проу вас есть собака, и постоянно пока- даем вовне. В частности, тот проект про расписание в салонах, о котором я уже говорил, один из самых эффекен этот рынок. Но пока ни одна ком- тивных с точки зрения оптимизации нашего собственного бизнеса.

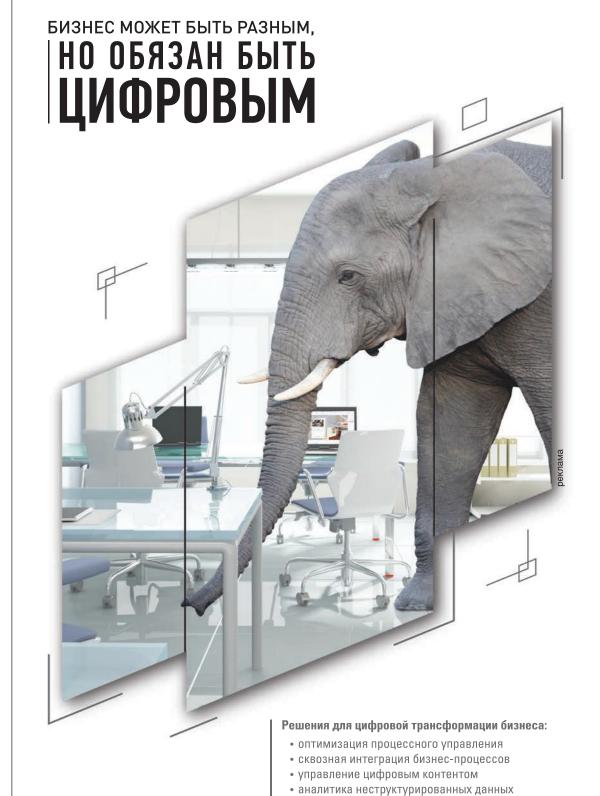
> Если говорить об остальных 20%, то, например, вместе с другими операторами связи мы продаем нологий Москвы агрегированную, обезличенную информацию о перемещении людей по городу. Эти данные учитываются при планировании развития транспортной инфраструктуры города. Проект суще-

— Вполне понятно, зачем это нуж- Да, в том числе мы добавляем но- но государственным службам. Есть ли какое-то взаимодействие с бизнесом по этому направлению?

— Конечно, аналогичные проекгаем определять, где лучше размереализованные проекты, напри- специфики бизнеса. Но нужно по- пользованием больших данных. Но мер с сетью частных клиник. Есть нимать, что этот рынок в России поготовые продукты для больших роз- ка только формируется. Мы раска-

циальных клиентов. Есть первые та разрабатываем решение с учетом повышать свою эффективность с исничных сетей. Под каждого клиен- чиваем его, показываем, как можно для таких технологий.

думаю, еще несколько лет уйдет на то, чтобы клиенты массово созрели







• роботизированные системы

