

Восемь подводных камней на пути к внедрению аналитики Больших данных

Решения для Больших данных

VERTICA

Аналитика Больших данных позволяет улучшить бизнес-процессы и операционную деятельность, повысить эффективность управления рисками и добиться дополнительной экономии средств.

Мир стремится использовать Большие данные

«Большие данные — золото XXI века. Если продолжить эту метафору, аналитика Больших данных является современным эквивалентом промывания золота»¹.

«По прогнозам IDC, к 2020 году организации, способные анализировать все необходимые данные и извлекать из них ценную информацию, повысят эффективность своей деятельности на

\$ 430 млрд

по сравнению с компаниями, меньше ориентированными на использование аналитики», —

Гил Пресс (Gil Press), Forbes⁴

Организации по всему миру тонут в океане информации, объем которой удваивается каждый год. Преимущества Больших данных существенны. Типичная компания из списка Fortune 1000, улучшившая доступ к данным на 10 %, может получить более 65 млн долл. дополнительной чистой прибыли². Для многих решение о создании программы внедрения и использования систем аналитики Больших данных — это неизбежность.

Однако совладать с Большими данными, аналитикой, автоматизацией и искусственным интеллектом не так-то просто. Скорее всего, на пути к Большим данным вы столкнетесь с серьезными препятствиями. Хотя интерес к аналитике Больших данных растет быстрыми темпами, только в 13 % проектов в этой области реализуется решение, которое можно внедрить в масштабе всей организации³.

Цель настоящего документа — помочь вам оставаться на верном пути. Наши рекомендации по продуктам и услугам позволят избежать восьми распространенных ошибок, связанных с программами аналитики Больших данных. Эти передовые методы были разработаны экспертами HPE на основе реального взаимодействия с заказчиками. Надеемся, что они приведут вас к цели. Мы также обсудим, как другие организации используют платформы аналитики Больших данных для решения широкого спектра бизнес-задач.

Восемь подводных камней

Если вы находитесь на начальных этапах планирования или внедряете аналитику Больших данных ускоренными темпами, следующие восемь советов помогут вам избежать ошибок.

1



Использование подхода, ориентированного на ИТ

Вы завершили исследование, провели анализ выгод и затрат и теперь переходите к проверке концепции проекта. Выполнив подробную техническую оценку, вы сузили круг возможных поставщиков. На этом этапе важно продолжить сотрудничество и обмен информацией с руководителями бизнес-подразделений, даже если вы в значительной степени сосредоточены на технических аспектах вашей программы внедрения и использования аналитики, ведь сценарии бизнес-использования могут меняться в соответствии с новыми требованиями бизнеса.

Эксперты The Data Warehouse Institute (TDWI) разработали метод оценки организаций на основе модели зрелости в области Больших данных. Эта оценка включает в себя 50 вопросов в пяти технических и бизнес-категориях: организация, инфраструктура, управление данными, аналитика и корпоративное управление. Указанные категории образуют комплексную картину, которая охватывает сотрудников и организацию, рассматривая их как важные элементы программы внедрения и использования аналитики Больших данных.

¹ «Что означают цифры и факты, стоящие за Большими данными? (What are the numbers, facts and figures behind Big Data?)», Бен Росси (Ben Rossi), Information Age, 4 марта 2016 г.

² «Большие данные: 20 ошеломляющих фактов, о которых должен знать каждый (Big Data: 20 Mind-Boggling Facts Everyone Must Read)», Бернард Марр (Bernard Marr), Forbes, 20 сентября 2015 г.

³ Руководство по достижению зрелости аналитики Больших данных (A Guide to Achieving Big Data Analytics Maturity), Ферн Халпер (Fern Halper) и Дэйв Стоддер (Dave Stodder), TDWI, июль 2016 г.

⁴ «6 прогнозов касательно аналитики Больших данных и когнитивных вычислений в 2016 году (6 Predictions For Big Data Analytics And Cognitive Computing In 2016)», Гил Пресс (Gil Press), Forbes, 15 декабря 2015 г.

Рекомендация. Рабочие встречи по вопросам проведения преобразований, проводимые совместно ИТ-службой и бизнес-подразделениями, позволят осознать преимущества бизнес-идей на основе аналитики и выгоды, которые можно было бы извлечь из лавинообразных потоков данных. Эффективная встреча дает возможность не только определить правильные процессы и инструменты, но и подготовить пилотный проект, который позволит проверить выбранную концепцию внедрения и использования аналитики. Эксперты помогут вашей организации реализовать изменения, выявить сотрудников, остающихся в стороне от процессов преобразований, и добиться единого понимания ценности данных и возможностей достижения поставленных целей с помощью внедряемых систем и методов аналитики.

«Во многих компаниях организационная структура все еще недостаточно продумана, а руководители не всегда понимают, что с помощью аналитики можно в корне изменить бизнес-процессы и модели поведения сотрудников. Как показало наше исследование, только 31 % компаний провели существенные преобразования в своих профильных подразделениях, добиваясь того, чтобы деятельность организации основывалась на данных»⁵.

2



Движение вперед при отсутствии нужных навыков и опыта

Аналитика Больших данных требует скоординированных усилий ИТ-службы, руководства компании, профильных подразделений и других служб, в том числе маркетинга. Вам понадобятся не только квалифицированные специалисты по данным, но и сотрудники ИТ-службы и бизнес-подразделений, способные сосредоточиться на аналитике Больших данных. Необходимо уделить внимание оценке как технических, так и бизнес-навыков.

Рекомендация. Новое поколение специалистов по данным использует для работы с предиктивной аналитикой такие инструменты, как Java, Python и R. Ваш партнер по внедрению поможет выполнить анализ несоответствия компетенций, чтобы понять, есть ли в вашей проектной команде специалисты с нужными навыками. Следует рассмотреть возможность проведения специализированного обучения в области Больших данных для сотрудников с техническими, управленческими и бизнес-компетенциями.

3



Планирование в условиях информационного вакуума

Обратитесь за советом к квалифицированным экспертам, способным в полной мере оценить ваши текущие возможности с позиций как технологий, так и бизнеса. Получите помощь от консультантов по стратегии, которые привнесут передовые методы в ваши процессы планирования и детальной проработки проектов аналитики Больших данных. Определите, может ли кто-либо из ваших внешних подрядчиков предложить хорошо зарекомендовавшее себя решение в области Больших данных. В условиях сжатых сроков такие услуги помогут осуществить планирование и внедрение аналитической платформы быстро и эффективно.

Рекомендация. Опытные консультанты по Большим данным помогут провести аудит ваших систем и определить, что необходимо для успеха вашего проекта. Крайне важно, чтобы ваша комплексная ИТ-стратегия поддерживала эффективное, безопасное управление данными и их глубокий анализ. В ходе специализированных консультативных встреч вы получите подробные схемы и планы действий по проведению преобразований, основанные на ваших конкретных требованиях. Также можно воспользоваться услугами по изучению данных вашей организации и выбору решений, наилучшим образом отвечающих вашим требованиям к производительности систем аналитики и удобству работы с ними.

4



Игнорирование вопросов безопасности

Не позволяйте себе забыть о безопасности данных. Наличие бреши в системе аналитики Больших данных чревато крупными неприятностями, особенно если речь идет о конфиденциальной информации клиентов. Репутация некоторых компаний серьезно пострадала, после того как о брешах стало известно общественности. Уже на ранних этапах планирования определите, как данные будут защищаться на протяжении всего их жизненного цикла: при хранении, передаче и анализе.

Рекомендация. Выберите правильные инструменты, чтобы преодолеть ограничения в области информационной безопасности, присутствие платформам аналитики Больших данных. Создайте центр передового опыта, чтобы в полной мере продемонстрировать бизнесу возможности ИТ перед их развертыванием. Кроме того, будьте внимательны к разрозненным элементам инфраструктуры, поскольку подходы к обеспечению безопасности в разных отделах могут быть несоординированными. Уровень безопасности данных определяется самым слабым звеном в инфраструктуре.

⁵ «Успех аналитики зависит от людей и организационных изменений (Analytics Success Depends on People, Organizational Change)», Крис Маццей (Chris Mazzei) и Гаурав Гупта (Gaurav Gupta), DataInformed, 16 июля 2016 г.

«Платформа Vertica помогает в реальном времени блокировать подозрительные действия и предоставлять ценную информацию о клиентах», —

Питер Хэнлон (Peter Hanlon), технический директор, Trader Media Group.

5



Недооценка сложности Больших данных

Мы живем в постоянно усложняющемся технологическом мире. Современные инфраструктуры данных получают информацию отовсюду: из бизнес-систем, таких как CRM и ERP, от датчиков, собирающих данные, которые генерируют машины, из Twitter и других социальных сетей, блогов и новостных лент, от газовых, электросетей и сетей сотовой связи — всего не перечислить. Даже просто хранение и управление этими объемами данных, собираемых из разных источников, представляет собой серьезную проблему, не говоря уже о глубокой аналитике.

Рекомендация. Многие организации решают проблему сложности, учреждая специальные должности, такие как главный архитектор данных. Эти руководители выбирают такие инструменты (коммерческие или с открытым кодом), которые смогут обеспечить эффективность всей цепочки работы с данными. Для успешного внедрения платформы аналитики Больших данных необходимы лучшие в своем классе решения, реализующие функции обработки потоков данных и средства их извлечения, преобразования и загрузки (ETL), создающие шины для обмена сообщениями о событиях, формирующие озера данных и использующие масштабируемые хранилища данных. Также важно выбрать удобное в использовании решение для бизнес-аналитики (BI). Чтобы ваша цепочка работы с данными была масштабируемой, надежной и не слишком дорогой, причем не только сегодня, но и в перспективе, необходимо понимать эволюцию этих технологий и их взаимосвязи.

6



Игнорирование вопросов безопасности

Вы уже вложили целое состояние в инфраструктуру бизнес-аналитики и хранилищ данных. Убедитесь в том, что выбранное вами решение может интегрироваться с имеющейся инфраструктурой и соответствует всем вашим требованиям независимо от того, где хранятся данные. Понимание возможностей платформы для работы с Большими данными может предрешить результат ваших усилий.

Рекомендация. Ваши технологии аналитики Больших данных должны помочь вам находить новые способы получения прибыли. Если ваша главная цель — монетизация данных или удержание клиентов, рассматривайте свою платформу как связующий компонент более широкой экосистемы Больших данных. Чтобы обеспечить максимальную отдачу от инвестиций, аналитическая платформа должна эффективно работать с существующей инфраструктурой — в частности, с инструментами преобразования данных и процессами ETL, а также с SQL-средствами визуализации.

7



Излишняя доверчивость

Рынок продуктов для работы с Большими данными — большой, шумный и полный назойливой рекламы. Приняв неверные решения в ходе подготовки проекта, вы можете поставить под удар свою программу внедрения аналитики. В ситуации повального увлечения Большими данными за шумихой, поднятой периферийными игроками, можно потерять из виду действительно ценные возможности. Эта сфера является динамичной и одновременно запутанной, даже для специалистов.

Рекомендация в отношении продуктов. Не принимайте на веру утверждения о том, что все ваши проблемы можно решить с помощью одного инструмента или платформы для работы с Большими данными. Оцените платформу аналитики и связанные с ней процессы с точки зрения гибкости, непрерывного исследования данных и интеграции. Обратите внимание на возможности развертывания и использования: желательно, чтобы платформа могла работать как в локальной среде, так и на базе Hadoop и в облаке. Наличие различных вариантов использования становится важным для получения желаемых результатов и максимальной окупаемости инвестиций.

8



Работа без учета передового опыта

После ввода аналитического решения в продуктивную эксплуатацию убедитесь в том, что ваши методы управления данными действительно являются лучшими. Стратегии в области Больших данных будут развиваться по мере роста требований вашего бизнеса. Компания, достигшая зрелости в сфере аналитики, будет искать ответы на вопросы: «Чьи это были данные? Чьи эти данные сейчас? Где они окажутся дальше? Как долго они будут использоваться?»

Рекомендация. Достигнув зрелости, ваша организация будет непрерывно использовать аналитику, решая новые бизнес-задачи. Глубоко продуманная программа корпоративного управления данными станет ключевым аспектом деятельности вашей организации. Руководство программой аналитики Больших данных в масштабах всей компании ляжет на проектный офис, курирующий инициативы в сфере Больших данных, и ваш управляющий комитет — они помогут вашей организации выполнить проекты в срок, в соответствии с бюджетом и целями бизнеса.

Перенимайте опыт успешных пользователей аналитики Больших данных

О новых областях применения аналитики Больших данных мы узнаем каждый день. Так, североамериканский телеком-оператор использует ее, чтобы улучшить обслуживание клиентов и избавиться от проблем с настройкой недавно приобретенных продуктов. Банк применяет аналитику для выявления закономерностей в покупательских привычках клиентов на всех этапах их жизненного цикла. Фармацевтическая компания с помощью аналитики стремится повысить предсказуемость продаж своей продукции крупным заказчикам. Сценарии использования аналитики в бизнесе становятся все более разнообразными, и предприятия уже взялись за оптимизацию управления программами аналитики Больших данных. Страховая компания из числа 50 ведущих игроков этого рынка вместо расширения имеющихся команд менеджеров вложила деньги в создание специальной группы по управлению изменениями, концентрирующей внимание на обучении персонала, внутренних коммуникациях и организационных преобразованиях с целью выработки эффективных программ в сфере аналитики⁶. В результате 77 % вовлеченных сотрудников отметили положительные сдвиги, помогающие их компании быть успешной в своей отрасли.

NASCAR

Национальная ассоциация гонок серийных автомобилей (NASCAR) искала способ отслеживать и анализировать освещение своей деятельности в различных СМИ. Компанию HPE она выбрала в качестве поставщика комплексного аналитического решения, включающего услуги, программное обеспечение и оборудование для создания современного центра вовлечения поклонников и представителей СМИ (Fan and Media Engagement Center, FMEC). По сути, FMEC представляет собой интеллектуальную платформу для сбора и анализа Больших данных на основе решения HPE Interactive Media Command Center (IMCC). В режиме реального времени она позволяет NASCAR выявлять, анализировать онлайн-обсуждения и реагировать на них. Такая возможность уникальна для индустрии развлечений. Например, во время гонок NASCAR осуществляет мониторинг и анализ около 14 000 онлайн-разговоров в минуту. Опираясь на долгий успешный опыт работы в отраслях коммуникаций,

«Безусловно, платформа HPE оказалась лучшим комплексным решением среди всех рассмотренных нами, — отмечает Стив Фелпс (Steve Phelps), исполнительный вице-президент и директор по маркетингу NASCAR. — А уровень консультационных услуг HPE значительно выше, чем у кого-либо другого».

⁶ «Успех аналитики зависит от людей и организационных изменений (Analytics Success Depends on People, Organizational Change)», Крис Маццей (Chris Mazzei) и Гаурав Гупта (Gaurav Gupta), DataInformed, 16 июля 2016 г.

СМИ и развлечений, HPE предложила NASCAR консультационные услуги, обеспечившие точное соответствие проекта функциональным требованиям заказчика и его своевременное выполнение.

Вы не одиноки в своем путешествии

Вот некоторые другие удивительные результаты, полученные заказчиками HPE с помощью аналитики Больших данных.

- Компания Intuit обрабатывает миллиарды транзакций, чтобы обеспечить быстрый персонализированный возврат налогов миллионам пользователей системы расчета налогов TurboTax.
- Организация Conservation International помогает ученым оценить влияние климатических, техногенных и агротехнических факторов, выполняя сравнительный анализ различных объектов и образцов по 86 миллионам записей в режиме, близком к реальному времени⁷.
- Корпорация Cerner на 6000 % повысила производительность системы для анализа эффективности клинических врачей по электронным медицинским записям (EMR), существенно улучшив качество медицинского обслуживания.
- Компания Supercell, разработчик игр, применяет аналитику для изучения в реальном времени поведения пользователей своих игр, тестирования продуктов и повышения уровня удовлетворенности 100 миллионов игроков ежедневно.
- Национальный комитет демократической партии США использует управляемый данными маркетинг и предиктивное моделирование, чтобы лучше понимать и прогнозировать поведение избирателей.
- Компания Guess ускорила подготовку важных отчетов в 90–400 раз, что позволило всем — дизайнерам, специалистам по закупкам и планированию производства, менеджерам торговых точек — лучше обслуживать клиентов.
- Фирма Criteo стимулирует продажи через Интернет с помощью встроенной в базы данных предиктивной аналитики, которая позволяет охватить 1,1 млрд интернет-пользователей, предлагая им рекламу, точно соответствующую их интересам.
- Специалисты по маркетингу оператора мобильной связи MTS India инициируют рекламные кампании, основываясь на проводимом в реальном времени анализе данных об удовлетворенности абонентов.
- Компания Tarjoу применяет аналитику, чтобы лучше понять влияние рекламы на поведение пользователей и повысить эффективность рекламных объявлений.
- Банк Finansbank использует аналитику для быстрого определения базовых характеристик и профилирования поведения пользователей. Выявленные аномалии помогают повысить безопасность и улучшить защиту от мошеннических операций.

«Программное обеспечение Vertica расширяет горизонты бизнеса», —

Кришан Кумар (Krishan Kumar), вице-президент по ИТ, Mzi Healthcare.

Как мы можем помочь вам

Как следует из наших рекомендаций, подводных камней при работе с аналитикой Больших данных можно избежать, используя целостный подход по превращению организации в гибкую, управляемую данными структуру. Для этого необходимы правильная инфраструктура и проверенный партнер с нужными навыками, знаниями и опытом. Мы поможем осуществить преобразования и получить выгоду от аналитики Больших данных без ущерба для уже сделанных инвестиций и существующих процессов. Ниже приведено краткое описание решений, обеспечивающих комплексный подход к аналитике Больших данных.

Аналитическая платформа Vertica

HPE Vertica обеспечивает высокую скорость, масштабируемость и простоту выполнения аналитических запросов с помощью распределенной колоночной архитектуры, использующей сжатие данных. HPE Vertica — это непревзойденная скорость (выполнение запросов в 50–1000 раз быстрее, чем в традиционных СУБД), масштабируемость до петабайтов (можно хранить на каждом из серверов в 10–30 раз больше данных, чем в традиционных СУБД), открытость и простота (можно использовать любые средства бизнес-анализа и ETL, а также Hadoop) — и всё это с гораздо более низкими затратами, чем при использовании традиционных решений для хранилищ данных⁸.

Таблица 1. Платформа аналитики Vertica

Предоставляет заказчикам следующие возможности:	Предлагает различные модели использования, в том числе:
<ul style="list-style-type: none"> • Полная поддержка стандартных и расширенных аналитических функций SQL 	<ul style="list-style-type: none"> • С локальным развертыванием (Vertica Enterprise [версии Premium и Express])
<ul style="list-style-type: none"> • Кластерный подход к хранению данных, обеспечивающий необходимую скорость выполнения поисковых и аналитических запросов 	<ul style="list-style-type: none"> • С развертыванием в облаке (Vertica on Amazon, применяется с нашей версией Amazon Machine Image [AMI])
<ul style="list-style-type: none"> • Эффективное сжатие данных, позволяющее снижать планку требований к оборудованию и хранилищу по сравнению с аналогичными решениями 	<ul style="list-style-type: none"> • По требованию (Vertica OnDemand)
<ul style="list-style-type: none"> • Гибкость и масштабируемость, дающие возможность повышать производительность по мере увеличения аналитической нагрузки 	<ul style="list-style-type: none"> • С хранением данных в Hadoop (Vertica SQL on Hadoop) — позволяет с легкостью анализировать данные, размещенные на узлах Hadoop
<ul style="list-style-type: none"> • Одновременное выполнение загрузки данных и запросов к ним 	
<ul style="list-style-type: none"> • Встроенные функции предиктивной аналитики, а также возможность использования библиотек с применением Python и R 	
<ul style="list-style-type: none"> • Снижение затрат и усилий на администрирование платформы и ее оптимизацию 	

Только Vertica предлагает столь широкий спектр режимов использования всех ключевых возможностей аналитической платформы, обеспечивая при этом лучшие в своем классе показатели производительности и масштабируемости. Решение Vertica безупречно работает в сложной экосистеме Больших данных, предлагая различные варианты их визуализации, интеграции и обработки, а также коннекторы для работы с данными всех форматов и из любых источников.

«С помощью Vertica, используя произвольные запросы, мы извлекаем ценную информацию из Больших данных с такой скоростью, о которой пользователи других систем могут только мечтать», —

Бен Уайт (Ben White), архитектор, TK Consulting.

⁷www.youtube.com/watch?v=ox-QtXq4mac

⁸www.techvalidate.com/tvid/B9F-BA0-073



Услуги Vertica professional services

Услуги Vertica professional services помогают заказчикам быстрее добиться бизнес-результатов, снизить затраты на разработку и обслуживание ИТ и получить конкурентные преимущества. Подразделение HPE Big Data Software Services выступает в качестве партнера на каждом этапе внедрения и использования аналитики Больших данных и предлагает широкий спектр услуг:

- Оценка и стратегический консалтинг, в том числе проведение рабочих встреч, планирование и подготовка «дорожных карт».
- Установка программного обеспечения, планирование необходимых мощностей, подготовка спецификаций оборудования, модели безопасности и общее управление.
- Услуги Data Source to Data Consumer Expert Services обеспечивают потребности заказчика в нужных компетенциях в нужное время и с прогнозируемыми затратами. Высококвалифицированные многопрофильные специалисты успешно справятся с интеграцией Больших данных и другими долгосрочными проектами в области консалтинга и интеграции.
- Услуги, предоставляемые после развертывания и ввода аналитических решений в эксплуатацию (Solution Management Services, SMS), охватывают следующие области:
 - реактивные услуги (единая в рамках всего решения точка контакта для управления инцидентами и проблемами);
 - эксплуатационные услуги (выполнение оперативных запросов и текущее обслуживание для обеспечения работоспособности решения);
 - расширенные услуги (непрерывное совершенствование решения путем улучшений, которые повышают его ценность и расширяют сферу использования).
- Услуга Health Check Service предлагает детальную оценку инфраструктуры для развертывания решения и включает в себя аудит с применением комплексного системного анализа операционной среды, а также подготовку рекомендаций по максимально полному извлечению выгоды из использования аналитической платформы.

Узнайте больше на сайте —

www.vertica.com

www.vertica.com/services/