**BIG DATA – общие проблемы**

Оглавление

[Как большие данные меняют парадигму бизнеса 1](#_Toc530238011)

[Есть данные: как big data помогает разобраться в проблемах бизнеса 5](#_Toc530238012)

[Люди не осознают масштаба катастрофы 7](#_Toc530238013)

[Большие данные как большие риски 12](#_Toc530238014)

[Big Data уже нельзя назвать инновацией 20](#_Toc530238015)

**[Вернуться в каталог сборников по менеджменту](http://учебники.информ2000.рф/management2/management3.shtml)**

[**Вернуться в электронную библиотеку по экономике, праву и экологии**](http://учебники.информ2000.рф/index.shtml)

[**НАПИСАНИЕ на ЗАКАЗ: дипломы и диссертации, курсовые и рефераты. Переводы с языков, он-лайн-консультации. Все отрасли знаний**](http://учебники.информ2000.рф/napisat-diplom.shtml)

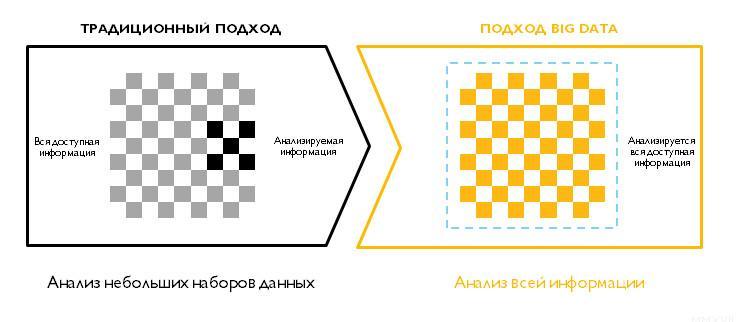
# Как большие данные меняют парадигму бизнеса

Искандер Шаймарданов из [Nectarin](http://www.nectarin.ru/) рассказывает о двух подходах к big data и намекает, как большие данные помогут даже небольшому бизнесу.

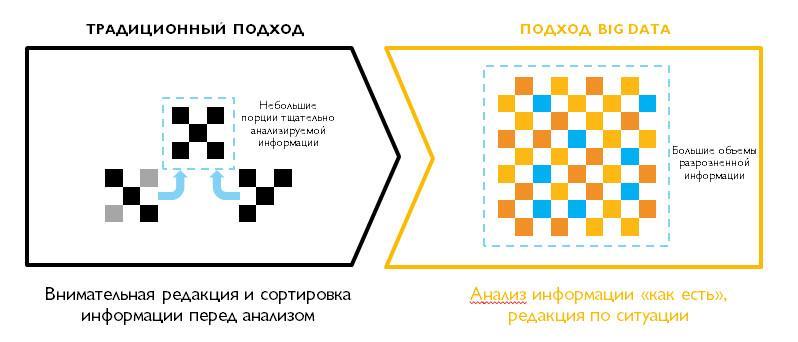
[Искандер Шаймарданов](https://www.cossa.ru/profile/?ID=130288), Nectarin 20 декабря 2017

**Традиционный подход VS большие данные**

Разберём основные преимущества использования big data над традиционным анализом небольших наборов данных.



При традиционном подходе анализируют только маленькие порции информации, тогда как big data позволяет исследовать весь доступный массив данных в поисках корреляций. Кроме того, смена подхода ведёт к постепенному снижению затрат на использование данных.





Упрощается процесс анализа данных. При традиционном подходе это возможно сделать только после того, как они собраны, обработаны и добавлены в хранилище. В случае с big data анализ данных происходит параллельно со сбором — в режиме реального времени.



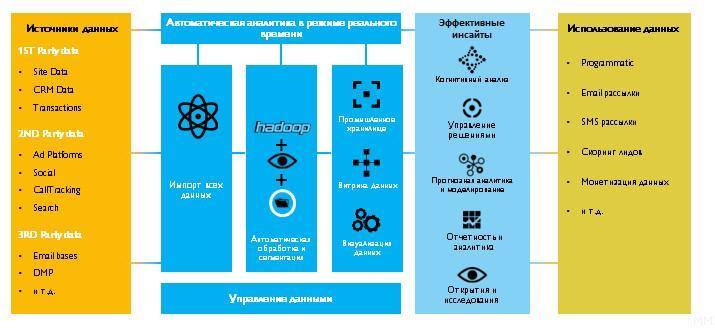
**Системы управления данными**

Теперь поговорим о двух основных системах управления данными: идеальной и оптимальной.

Идеальная система

Массив данных импортируется на клиентский сервер (хранилище). Затем сырые данные обрабатывают и визуализируют с помощью[фреймворка Hadoop](https://ru.wikipedia.org/wiki/Hadoop) и его библиотек.

Интерфейс визуализации пишется индивидуально, с отображением именно той информации, которая необходима. В дальнейшем любые полученные данные на основе эффективных инсайтов можно использовать в рекламной кампании.



Драйверы внедрения идеальной системы:

* визуализация данных в режиме реального времени;
* возможность работать с действительно большими массивами данных;
* автоматическое управление данными;
* прогнозная аналитика;
* более точные инсайты.

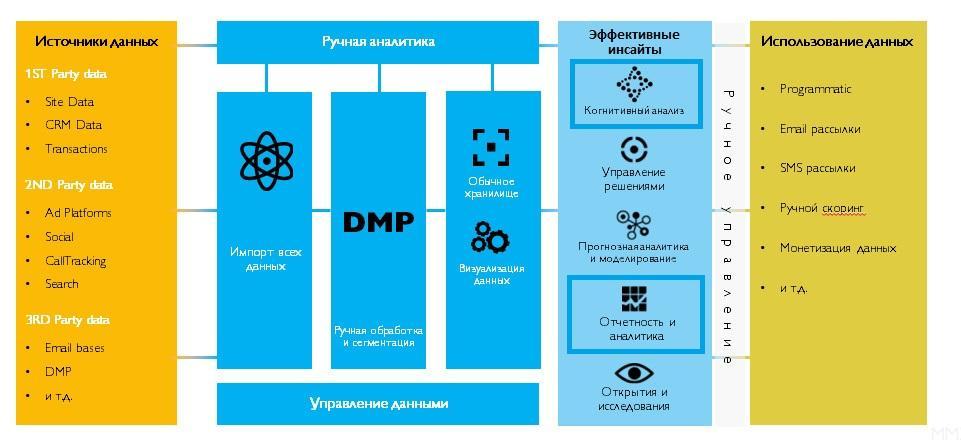
Барьеры идеальной системы:

* высокая стоимость внедрения;
* проблема выбора обрабатываемых данных;
* нехватка специалистов (аналитик бизнес-модели, аналитик структур данных, администратор хранилища данных);
* устаревшие бизнес-процессы или внутренние регламенты;
* передача данных в третьи руки.

Оптимальная система

Данные поступают на сторонний сервер платформы управления данными (DMP). В рамках этой DMP доступно хранение данных на оплаченный период, а также визуализация по заранее предустановленным шаблонам.

Обработанные данные в дальнейшем можно вручную выгружать в сторонние сервисы для настройки рекламных кампаний.



Драйверы внедрения оптимальной системы:

* невысокая стоимость внедрения;
* скорость внедрения;
* контроль данных, упрощённая система работы и безопасность;
* грамотная сегментация, глубокое понимание целевой аудитории;
* анализ аудитории.

Барьеры оптимальной системы:

* ручное управление данными;
* меньше возможностей для обработки и интерпретации;
* невозможно работать с большими массивами данных;
* ниже точность выводов.

**Примеры использования big data**

Кейс PropertyGuru

[PropertyGuru](https://www.propertyguru.com.sg/) развивается в двух направлениях: email-маркетинг на клиентов, ищущих дом, и на риэлтеров, которые выступают экспертами в этом вопросе.

PropertyGuru мотивирует обе стороны с помощью рыночной аналитики и другой полезной информации. До недавних пор PropertyGuru полагался на ряд не связанных между собой маркетинговых активностей и ситуационных кампаний.

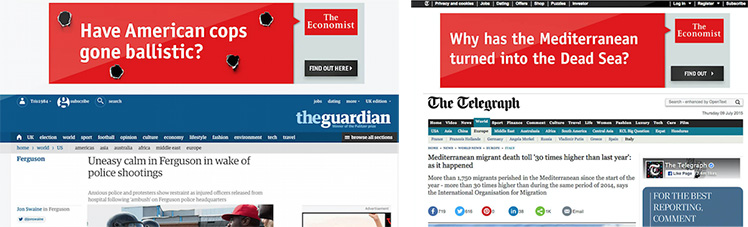
Для повышения эффективности кампаний PropertyGuru использовал Oracle DMP. Система создания динамического контента позволяет PropertyGuru рассылать агентами информацию со сравнением их прогресса и рекомендациями по его улучшению, включая данные о том, сколько объектов размещено, лиды, которые они собрали, сведения о ресурсах, которые помогут повысить тариф. Также агентам рассылается важная для ежедневной работы информация, например, о рыночных новостях.

[Результат](https://www.oracle.com/marketingcloud/customers/success-stories/propertyguru.html): автоматизированные рассылки позволили значительно повысить open rate (до 40%) и click through rate (до 4%).

Кейс The Economist

[The Economist](https://www.economist.com/) для решения стратегической задачи — привлечения платных подписчиков — сфокусировался на анализе аудитории и её интересов.

Сопоставляя cookie, информацию о подписчиках и другую аналитику, команда проекта получила лучшее представление о сегментах читателей. Это в свою очередь позволило настроить рекламную кампанию на основе технологии look-alike — таргетированную на потенциальных пользователей и подписчиков.

[](https://www.cossa.ru/upload/medialibrary/b2b/image9.jpg)

The Economist запустил 12-недельную кампанию с использованием инструментов дисплейной и контекстной рекламы. Вопросительные заголовки объявлений обыгрывали актуальные темы дня и в реальном времени таргетировались на те сегменты, для которых именно эта тема может быть наиболее актуальна.

При клике на объявление пользователь переходил на посадочную страницу с возможностью подписки. The Economist создал 60 объявлений с целью привлечь 650 000 потенциальных подписчиков. Цели достигли за 9 дней. При этом CPA снизился на 50%.

[В результате кампании](https://www.oracle.com/marketingcloud/customers/success-stories/the-economist.html) удалось привлечь 3,6 миллиона потенциальных подписчиков и более 10 000 реальных (платных). А показатель ROI превысил 10:1 к запланированному.

# Есть данные: как big data помогает разобраться в проблемах бизнеса

**19 июня 2018 Ана Брант**

Чтобы предлагать гостям отелей класса ультра-люкс Dorchester Collection самое лучшее и сделать маркетинг качественнее, мы прибегаем к помощи больших данных и аналитики. Наш аналитический инструмент Metis анализирует данные онлайн-обзоров и социальных сетей, чтобы выявить проблемы и возможности. Я работаю директором по гостеприимству и инновациям в Dorchester Collection и обнаружила, что зачастую данные позволяют определить только саму проблему, но не причину ее появления или способ устранения. Для этого нужно участие человека.

Например, в прошлом году мы с помощью Metis оценили впечатления гостей после пребывания в парижских отелях класса люкс. Оказалось, что гости были не очень лояльны к нашим отелям Le Meurice и Hotel Plaza Athénée и к отелям конкурентов. Как показывает анализ Metis, гости не видят большой разницы между пятизвездочными отелями Парижа. Разные отели они посещают, просто чтобы попробовать что-то новое. После того, как анализ Metis показал недостаточную лояльность клиентов, перед нами встала задача выяснить причины и понять, что делать дальше.

**Наблюдения и исследование в Париже**

Мы начали с изучения рынка. По версии Forbes в Париже есть десять пятизвездочных отелей (больше их только в Макао, где в этом году еще два отеля [попали в рейтинг](https://www.prnewswire.com/news-releases/forbes-travel-guide-announces-2018-star-rating-awards-300600578.html)). Во всех пятизвездочных отелях Парижа есть рестораны, отмеченные звездами Мишлен. Все они предлагают роскошные приветственные подарки при заезде (шампанское, шоколад), номера в них приблизительно одного размера и стоимость примерно одинакова.

Не удивительно, что для клиентов они равноценны. Нам предстояло сделать так, чтобы наши отели отличались.

Сотрудники Plaza Athénée наблюдали за тем, какую одежду предпочитают гости отеля и какие магазины они посещают. Конечно, многие из них или были представителями мира высокой моды, или четко ей следовали, поэтому мы решили позиционировать Plaza Athénée как отель haute couture, где в 1947 году Кристиан Диор представил публике свою первую коллекцию. В отеле открылся спа-центр Dior, а в ванных комнатах разместили косметику Guerlain. Был создан зал для торжеств Le Salon Haute Couture, а во время недели моды в ресторане был организован сервис в стиле Dior. Для отеля выбрали слоган «Важно не то, какая на вас одежда. Главное — куда вы в ней отправитесь».

В свою очередь, Le Meurice известен тем, что здесь часто останавливались художники (здесь жил Сальвадор Дали во время своего пребывания в Париже). Сотрудники отеля заметили, что многие гости приезжают, чтобы посетить галереи и музеи, поэтому было решено, что Le Meurice будет покровительствовать современному искусству. Помимо прочего, за две недели до Международной ярмарки современного искусства FIAC была учреждена премия в 20 тысяч евро для перспективных художников, а классические предметы искусства, характерные для парижских пятизвездочных отелей, заменили на произведения современного искусства. Один из ресторанов переименовали в Le Dali.

Пока прошло не так много времени, но в Dorchester Collection уже увидели рост лояльности клиентов обоих отелей по сравнению с прошлым годом. Неожиданным преимуществом стало полноценное освещение изменений в прессе и социальных медиа, что привлекло к отелю внимание большего количества представителей местного бизнеса. Для бренда всегда полезно обзавестись поддержкой местных жителей (особенно если это парижане).

Итак, Metis указал нам на проблемы с лояльностью, но определить, что именно нужно сделать для их решения, уже было задачей для людей.

**Как мы наладили свадебный сервис в Лос-Анджелесе**

Мы также применили Metis для анализа свадебного сервиса. За несколько лет число свадеб, проведенных в отелях Dorchester Collection в Лос-Анджелесе (Hotel Bel-Air и The Beverly Hills Hotel), возросло. Однако доходы от этого не повысились. Нам нужно было узнать, почему.

По данным Metis, отзывы от гостей свадеб были скорее нейтральными, чем положительными. Жених и невеста, что неудивительно, озабочены тем, чтобы все держать под контролем, а семья и гости озабочены тем, чтобы помочь жениху и невесте. К тому же семье сотрудники отеля уделяют не так много внимания. Как только мы узнали о проблеме, настало время изучить процесс от начала до конца — от приглашений до благодарственных писем.

Команда по инновациям занялась поиском в Google приглашений на свадьбу в The Beverly Hills Hotel. Оказалось, что многие пары сами копировали классическое изображение бананового листа на приглашениях, добавляя себе работы. Команда составила перечень предпочтительных полиграфических компаний, предоставив им наш логотип и дизайнерские решения, после чего посетителей, готовящих приглашения на свадьбу, стали направлять в эти компании.

Кроме того, команде удалось выяснить у наших организаторов мероприятий, что многие невесты бронируют отель до того, как отправляются на поиски «самого главного платья», а это занятие весьма изматывающее, ведь многие тратят на наряд десятки тысяч долларов. Мы стали сотрудничать с магазинами платьев Rodeo Drive и направлять невест туда: магазин дает покупательницам скидку, мы получаем комиссию, и так появляется новый источник дохода.

Чтобы повысить качество обслуживания участников свадебного торжества, организаторам поручили составлять семейное древо, чтобы было понятно, кого можно привлекать для помощи и кому уделять особое внимание.

Наконец, нашу свадебную программу стало проще найти на сайте, и свадебные услуги стали отличительной чертой отеля на рынке.

Благодаря этим изменениям доходы отелей возросли, а впечатления гостей от свадебных мероприятий улучшились.

**Сочетание машинного обучения и человеческих наблюдений и суждений**

Использование больших данных и аналитики без человеческих суждений способно сделать неожиданные выводы: люди скучают с шампанским и шоколадом, свадьбы вызывают раздражение у женихов и невест, их семей и друзей. Информационные инструменты не всегда позволяют определить причину того или иного отношения людей или найти способ решения проблем. Только человек может взглянуть на проблему, понять ее суть и найти решение, полагаясь на собственные чувства и разум.

**Об авторе.** Ана Брант — директор по международному гостеприимству и инновациям Dorchester Collection в Лондоне. Ранее занимала должности руководителя по качеству в The New York Palace и регионального директора по качеству в The Beverly Hills Hotel и Hotel Bel-Air. В начале карьеры работала в The Ritz-Carlton Hotel. Выступала с докладами в Гарвардском университете, на конференции лауреатов премии Болдриджа и на саммите по исследованиям в сфере гостеприимства Корнеллского университета в 2014 году.

# Люди не осознают масштаба катастрофы

**20 августа 2018 Анна Натитник**

*В эпоху высоких технологий человеческое общение становится важнейшим элементом, а взаимодействие с потребителями — краеугольным камнем маркетинга. Заместитель декана по исследованиям и профессор маркетинга Бизнес-школы им. Саида Оксфордского университета Эндрю Стивен на конференции Teradata Universe 2018 рассказал старшему редактору «HBR — Россия» о том, что нового компании узнают о клиентах и как перестраивают свою маркетинговую стратегию.*

**HBR — Россия: Какие характеристики маркетинга вы назвали бы основными в цифровую эпоху?**

**Стивен:** Сегодня все чаще возникают проблемы с защитой персональных данных, поэтому доверие становится одним из существенных элементов маркетинга. Но поскольку роботу, в отличие от человека, доверять нельзя, человеческое общение и межличностные связи в цифровой среде приобретают все большее значение.

**Человеческое общение — это личное взаимодействие с потребителями?**

Не обязательно личное — общаться можно и при помощи технологий. Но общение — это еще не все. Компаниям важно показать, что их продукты и услуги в точности отвечают нуждам и желаниям потребителей. Раньше считалось, что главное для бизнеса — создавать качественные продукты, потом — что важно строить сильные бренды, затем — что необходимо развивать технологии. Сейчас наступает новый этап: организации понимают, что важны и продукты, и бренды, и технологии — но, кроме того, еще и потребители. Все буквально помешались на клиентоориентированности. Проявлять заботу о клиентах уже недостаточно — их нужно хорошо знать. Успешные компании давно научились предвосхищать запросы потребителей. Теперь, благодаря технологиям, этим навыком могут овладеть все.

Поскольку будущее маркетинга связано с необходимостью взаимодействовать с людьми, особое значение приобретет персонализация — не только продуктов, но и опыта, общения.

**Как в связи с диджитализацией ­меняется структура компаний?**

В организационной структуре происходят сдвиги, традиционная иерархическая система становится более плоской. Такие перемены связаны и с внедрением технологий, которые значительно ускоряют все процессы, и с необходимостью переходить на гибкие методы управления. Поскольку решения все чаще приходится принимать оперативнее, работу выполнять — быстрее, возникает потребность в сокращении уровней иерархии.

Трансформируется структура команд, в том числе ­маркетинговых. Обычно маркетингом ведают бренд-менеджер, менеджер по продукту, специалист по потребительским предпочтениям и т. д. Сегодня в такие команды вводят специалистов по обработке данных, а также аналитиков. Если традиционно аналитикой занимался специальный отдел или сторонние организации, то теперь аналитики должны работать рука об руку с сотрудниками, ответственными за бренд, продукт, розницу. Так что на смену специализированным командам приходят межфункциональные.

Меняется и подход к оценке персонала. Компании понимают, что говорить об индивидуальном вкладе сотрудника, который работает в коллективе, неправильно, — нужно оценивать всю команду и награждать тоже всех. Бонусы становятся коллективными, а не индивидуальными.

**Должны ли современные маркетологи обладать навыками анализа данных?**

Конечно, не каждый должен быть профессиональным аналитиком данных, но понимать, в чем состоит работа этих специалистов, должны все. В противном случае маркетолог не справится со своей задачей — быть связующим звеном между аналитиками, с одной стороны, и продуктами и потребителями — с другой.

К сожалению, сегодня маркетологи, даже молодые, как правило, плохо разбираются в аналитике. Поэтому задача компаний — нанимать технически подкованных сотрудников, с одной стороны, и обучать уже имеющихся, с другой.

**Что нового бизнес может узнать о потребителях благодаря технологиям?**

Приведу пример. Компания L’Oréal традиционно распространяет продукцию через розничные магазины и получает от них сведения только об объеме продаж. Несколько лет назад она осо­знала, что плохо знает своих клиентов, и выпустила приложение, благодаря которому потребители могли «примерять» продукты — помаду, тени, румяна, — и экспериментировать с разными образами.

Получив массу информации о цветах и оттенках кожи, зафиксировав самые незначительные нюансы, L’Oréal собрала большую базу данных. Проанализировав ее, она поняла, какой путь проходят люди перед покупкой того или иного товара. В результате компания смогла создавать продукты, которые в большей степени отвечают спросу, — в частности, кастомизировала декоративную косметику Lancome. (Подробнее об этом приложении см. в статье Д. Эдельмана и М. Сингера [«Взять клиента под опеку и довести до покупки»](https://hbr-russia.ru/marketing/marketingovaya-strategiya/a16829), «HBR — Россия», декабрь 2015 г.).

Собранную таким образом информацию можно использовать с разными целями — не только в маркетинге, но и, например, в производстве. Классический анализ подобной глубины проникновения и понимания не дает.

**Как и почему сейчас трансформируется поведение потребителей?**

Последнее десятилетие оно меняется под влиянием соцсетей. Сегодня большинство людей делится своим опытом именно там. Интернет существенно облегчил взаимодействие потребителей и обмен информацией о товарах и услугах. Казалось бы: чем это отличается от сарафанного радио? Когда я рассказываю в Facebook, что был в таком-то ресторане, я не пишу: «Рекомендую туда сходить», так что мой пост — не отзыв в чистом виде. Это вроде обычного устного рассказа, но от лица клиента: я сообщаю о своем опыте потребления.

В последнее время влияние соцсетей выражается в том, что люди начинают задумываться, какую информацию публиковать, а какую нет: проблема приватности, или неразглашения личной информации, выходит на первый план. Посмотрим, что будет дальше.

**Понятие приватности в последнее время тоже существенно изменилось.**

Безусловно. Люди потихоньку начинают понимать, как работает таргетирование рекламы, как используются их персональные данные, как бизнес узнает об их предпочтениях, интересах, поведении. До недавних пор никто об этом даже не думал. (Любопытно, что, как показала встреча сенаторов США с Марком Цукербергом, Сенат тоже не имел об этом представления.) Недавние скандалы, например с участием Cambridge Analytica и Facebook, многих заставили насторожиться.

В мае в Евросоюзе вступил в силу Общий регламент по защите данных, который дает гражданам инструменты контроля над своими личными данными. Люди получили право интересоваться, кто и как использует сведения о них. Регламент имеет экстерриториальное действие, то есть соблюдать его должны все компании мира, обрабатывающие персональные данные резидентов и граждан ЕС. Думаю, это оздоровит взаимоотношения потребителей и бизнеса.

С человеческой природой, однако, бороться непросто, так что пользователи продолжат раскрывать о себе слишком много информации. Но, по крайней мере, они будут понимать, к чему это может привести.

**Чего именно люди не понимают, когда речь идет о сборе и использовании их данных?**

Думаю, люди догадываются: все, что они публикуют, отслеживается, анализируется и используется для рекламы. Однако далеко не все понимают, что платформы, на которых есть реклама (Facebook, Google и другие), знают, какие сайты посещают пользователи, получают данные о геолокации с их телефона, видят, в какие магазины они заходят. Люди не осознают масштаба «катастрофы», не подозревают, что о них собирают даже незначительные данные, не отдают себе отчета в том, на что соглашаются, нажимая кнопку «Да» в пользовательском соглашении.

Обратите внимание, насколько точно бьет в цель реклама Facebook. У нее потрясающие механизмы таргетирования с невероятным уровнем детализации. Они учитывают не только демографические показатели и местоположение, но и интересы пользователей, их потребности и желания. Откуда Facebook знает, что вы, например, подыскиваете себе машину? Вы посещаете сайт BMW и рассматриваете определенные модели, потом изучаете страницы, связанные с BMW, в Facebook, затем идете в офис дилера или в автосалон. Компания обо всем этом узнает: о посещении сайта с помощью «Пикселя» — инструмента, встроенного во многие платформы, который передает Facebook информацию о вашей активности; о походе в салон — благодаря включенной в вашем телефоне геолокации. С помощью несложных алгоритмов все это анализируется, сопоставляется, и вот уже Facebook знает, что вы ищете машину. Компания не продает ваши сведения рекламодателям, не передает им списки имен и электронных адресов потенциальных покупателей. Она предоставляет рекламодателям площадку и отбирает для них целевую аудиторию, опираясь на собранные данные.

Думаю, об этом люди, не связанные с цифровой индустрией, не имеют представления. Если им все объяснить, они, скорее всего, забеспокоятся.

**Узнав, что компании за ними следят, клиенты наверняка перестанут им доверять. Как этого избежать?**

Бизнес должен показать, что ответственно относится к персональным данным. Для этого ему нужно, во-первых, четко сформулировать свое понимание «ответственности», то есть заявить: мы не станем продавать ваши данные, будем их защищать и не допустим к ним хакеров. Во-вторых, нужно объяснить клиентам: мы собираем сведения не просто так, а для вашей пользы; благодаря полученной информации мы сможем повысить качество товаров и услуг, показывать вам только интересный контент, делать только важную рассылку.

За счет этого компания сможет построить не только доверительные отношения с потребителями, но и сильный бренд.

**И все же не все готовы мириться с утечкой своих данных. Можно ли в наше время оставаться невидимкой?**

Не думаю. Конечно, можно постараться оставлять минимум цифровых следов, но факт остается фактом: мы все «под колпаком». Даже если вы уйдете из сети, избавитесь от мобильного телефона и начнете пользоваться только наличными, вы все равно попадете в поле зрение камер — скажем, в магазине. Понятно, что не во всех странах видео с ­камер обрабатывается и используется, но во многих государствах оно прогоняется через систему распознавания лиц. Так что право на тайну частной жизни есть у всех, но возможности в полной мере его реализовать нет почти ни у кого.

С другой стороны, очевидно, что с нашими данными зачастую обходятся слишком вольно и это нарушает этические нормы. Если общество перестанет заострять на этом внимание, ничего хорошего ждать не стоит. Это опасно в том числе для демократии.

**Что вы думаете о праве государства следить за гражданами?**

Правительства всегда будут выхо­дить за границы дозволенного — не обязательно с дурными намерениями, а, например, из соображений безопасности. Конечно, мы все против тотальной слежки, но в наше время технологии достигли такого уровня, что организовать слежку за всеми гражданами проще простого.

Я считаю, что угроза исходит не только от правительств, но и от крупных компаний, которые обладают доступом к огромным массивам данных, — Google, Facebook, Amazon и др. Голосовые сервисы вроде Alexa Voice Service или Siri постоянно нас слушают. И если Apple заявляет, что все голосовые данные обрабатываются на самом устройстве и не хранятся в облаке, то Amazon про Echo и Alexa этого не говорит: все полученные ими данные можно восстановить. Я не утверждаю, что эти организации все записывают и могут обратить против нас, — но кто знает?

Я убежден: даже если государство введет законы, регулирующие обращение с персональными данными, все равно найдутся люди, которые будут их нарушать — даже под угрозой сурового наказания.

**Цифровые устройства достигли такой стадии развития, что способны принимать решения за людей. Можем ли мы в этом на них полагаться?**

В какой-то степени мы уже на них полагаемся — вспомните хотя бы рекомендательные сервисы, в которых специальные алгоритмы вычисляют, какие фильмы, концерты, книги могли бы нам понравиться. Например, Netflix рекомендует следующий фильм для просмотра на основании наших предпочтений.

**Но сейчас уже говорят об устройствах, которым вы будете поручать более сложные задачи — например, самостоятельно выбрать и купить подарок жене.**

Сегодня мы спокойно позволяем цифровым устройствам подбирать для нас разные варианты, а потом сами делаем окончательный выбор. То, о чем вы говорите, — следующий шаг. Он вполне реален. Если я, например, уже много раз принимал подобные решения (скажем, выбирал подарок жене), то алгоритм выучил, что мне нравится, и наверняка сможет сделать за меня правильный выбор. Почему бы это не автоматизировать?

Кстати, предтечи таких алгоритмов уже существуют. У Amazon, например, во многих странах есть опция «Подпишись и сохрани». Предположим, вы постоянно покупаете кошачий корм определенной марки. Amazon говорит: «Зачем делать это вручную? Мы дадим вам 5-процентную скидку, если вы оформите подписку. Мы будем доставлять вам этот товар с выбранной вами регулярностью, пока не передумаете». Это выглядит как преданность определенной марке, но на деле отражает стремление упростить процесс закупок: человек делает выбор один раз, а затем автоматически его повторяет.

Эта тенденция порождает новые типы потребительского поведения и меняет подход к маркетингу. Бизнесу придется искать дополнительные способы привлечения клиентов. Вместо того чтобы убеждать людей купить продукт компании или вновь воспользоваться ее услугами, маркетолог должен будет приложить все силы, чтобы внести этот продукт или услугу в базу рекомендательной системы.

**Если все решения станут принимать обучаемые алгоритмы, то от людей будет зависеть все меньше?**

Некоторые технологии действительно сужают наше поле зрения. Посмотрите, например, как люди потребляют новости. Они читают их в соцсетях или в специальных приложениях-агрегаторах, которые учитывают интересы и политические взгляды пользователя и выдают ему только релевантные новости. То есть люди окружают себя стеной, препятствующей проникновению извне любой не интересующей их информации. Это значительно поляризует общество. Например, либерал, полагающийся на алгоритмы поиска новостей, читает только либеральные статьи, консерватор — только консервативные. По большому счету они не знакомы с точкой зрения оппонентов, и это нездорово.

Если рассмотреть эту проблему в контексте маркетинга, окажется, что потребитель всегда получает информацию только о тех продуктах, которыми он уже пользуется и о которых знает. Значит, он никогда не откроет для себя новые товары.

Разработчики алгоритмов наконец поняли, что проблему необходимо решать. Они осознали, что у людей должен быть разнообразный выбор, и перепрограммируют алгоритмы так, чтобы добавить в их работу элемент случайности.

**Останется ли в цифровом мире место для традиционных физических магазинов или все будет автоматизировано и перейдет в онлайн?**

Думаю, традиционная торговля никуда не денется, но ее характеристики изменятся. Потребители любят офлайновые магазины, даже если поход туда кажется нерациональным. Да, они могут заказать продукты через интернет: это экономит время, — но все равно пойдут в супермаркет, чтобы посмотреть, что там появилось, найти что-нибудь новое. Физический ритейл дает возможность погрузиться в особую среду, получить дополнительный опыт. Он поддерживает в людях желание экспериментировать и что-то для себя открывать. Я, например, люблю готовить, и меня традиционный магазин вдохновляет, дает свежие идеи; я нахожу неожиданные ингредиенты, которые мне бы и в голову не пришло искать через интернет.

Кроме того, некоторые магазины — это «храм бренда». Классический пример — магазины Apple. Обратите внимание: там всегда много народу, и кажется, все что-то покупают. Это при том, что у Apple отлично развито онлайн-направление. То есть в физический магазин люди идут ради опыта, чтобы пощупать товары, подер­жать их в руках, попробовать на них поработать, получить помощь и техническую поддержку, на­учиться чему-нибудь в «Баре гениев». Конечно, что-то из этого можно сделать онлайн, и в будущем, вероятно, возможности расширятся за счет виртуальной реальности. Но потребность в тактильных ощущениях по-прежнему останется.

Я убежден, что будущее ритейла не связано с конкуренцией онлайн- и офлайн-магазинов. Будущее — за омниканальностью: ритейлер должен обеспечить потребителю возможность совершать покупки в любом удобном для него месте в любое удобное время.

Конечно, количество физических магазинов существенно сократится. Люди будут реже их посещать: если они однажды сделали выбор и знают, что им нужно, то рассматривать и трогать товары не обязательно. Уже сейчас во многих странах торговые центры пустуют. Часть из них, полагаю, ликвидируют, а часть изменит свою функцию — туда будут ходить, скорее, не за покупками, а чтобы развлечься.

# Большие данные как большие риски

Все и повсюду массово и с пользой накапливают гигантские базы данных. Беда в том, что никто и нигде не знает, как эти горы данных защищать от злоупотреблений  
  
Проблемы разные – суть одна  
  
В ленту текущих новостей из разных концов планеты недавно попали три сообщения, рассказывающие о существенно разных, казалось бы, проблемах у людей и властей в связи с особенностями новейших инфотехнологий. Единственное, на первый взгляд, что их объединяет, – это общая тема рисков и угроз, с которыми сталкивается общество по мере все более масштабного освоения Больших данных.  
  
Но вот если с этими неприятностями начать разбираться чуть внимательнее, то довольно быстро становится ясно, что на самом деле речь везде идет о разных сторонах одной и той же проблемы. Такого рода серьезные вещи обычно принято именовать «системными дефектами конструкции». Специалисты-профессионалы об этой беде, конечно же, отлично знают. Вот только как это дело эффективно лечить – вопрос более чем дискуссионный, поскольку варианты ответов сильно зависят от того, кому они адресованы.  
  
Иначе говоря, владельцы баз с Большими данными трактуют безопасность системы одним образом; профессионалы ИТ-сектора, занимающиеся разработкой и анализом защиты подобных систем, видят проблемы по-другому; ну а те люди, чьи персональные данные накапливаются и обрабатываются в этих системах в гигантских количествах, представляют себе заботу об их личной, чувствительной к разглашениям информации существенно иначе — по-третьему.  
  
Проиллюстрировать суть возникающих здесь проблем и коллизий, ясное дело, удобнее всего на живых примерах. Самое время процитировать сообщения из ленты текущих новостей.  
  
В Индии члены Верховного суда своим единогласным решением постановили, что персональная приватность, или право на тайну личной жизни, является фундаментальным правом каждого человека. Теперь вердикт высшей в стране судебной инстанции не только должен оказать существенное влияние на то, каким образом корпорации обращаются с персональными данными граждан, — он уже нанес серьезный удар по намерениям нынешнего правительства во главе с Нарендрой Моди. Это самое правительство пыталось доказать, что право на приватность не является фундаментальным правом человека, защищаемым конституцией, а потому все граждане якобы обязаны предоставлять свои отпечатки пальцев, снимки радужки глаза и прочую биометрию в общенациональную биометрическую базу данных, присваивающую на этой основе каждому индийцу уникальный номер-идентификатор.  
  
Новость номер два: в Швеции разгорелся серьезный политический скандал вокруг национального Транспортного агентства. Высшее руководство этого правительственного ведомства решило сэкономить на расходах, воспользовавшись столь модным нынче аутсорсингом – перенеся свои базы данных в облачное хранилище компании IBM, территориально расположенное в другой стране. Вообще говоря, законы Швеции подобные вещи не запрещают. Вот только именно в этом конкретном случае базы данных Агентства содержат немало такой информации, которая может классифицироваться как государственная тайна (например, персональные данные всех водителей транспортных средств в вооруженных силах и спецслужбах; данные о всех гражданах, сменивших имена по государственной программе защиты свидетелей; ну и прочие подобные сведения весьма деликатного характера). Однако позаботиться о том, чтобы доступ ко всем этим сведениям, размещенным в зарубежном хранилище, имели исключительно люди с соответствующим образом оформленным допуском к секретам, руководство Агентства почему-то не сочло необходимым. Попутно оно наделало и других ошибок, скомпрометировавших доверенные им массивы чувствительных к разглашению данных.  
  
Из США пришла новость номер три — о серьезном конфликте, назревшем между крупными полицейскими управлениями мегаполисов (Лос-Анджелеса и Нью-Йорка), с одной стороны, и знаменитой датамайнинговой компанией Palantir — с другой. Поначалу мощные системы Palantir были разработаны для помощи в аналитической работе разведывательного сообщества и спецслужб типа ФБР – чтобы увязывать, синтезировать и визуализировать большие массивы разнородных данных из множества источников.  
  
А еще через несколько лет, когда системы анализа БД от Palantir начали охватывать не только особо скрытные федеральные спецслужбы, но и региональные управления полиции, попутно стали появляться и сигналы о существенных проблемах с этим инструментарием. В частности, ныне стало известно, что системы Palantir не обеспечивают эффективное «секционирование» (compartmentation) информации, то есть не способны надлежащим образом контролировать выдачу секретных сведений разным пользователям – будь то внутри одного учреждения или даже при объединении данных от разных ведомств. По этой причине – в конкретных условиях полиции – тайный осведомитель отдела по борьбе с организованной преступностью становится видимым и известным, к примеру, для сотрудников дорожной полиции и для полицейских совсем других подразделений. Последствия таких утечек, как несложно догадаться, могут оказываться весьма трагическими.  
  
Из трех приведенных примеров уже понятно, наверное, что правильная работа с Большими данными – это дело весьма и весьма непростое. А грубейшие ошибки здесь продолжают регулярно совершать даже те профессионалы, которых принято считать лидерами отрасли. Поэтому для более глубокого понимания проблем каждую из перечисленных новостей имеет смысл разобрать чуть подробнее – за конкретными национальными особенностями во всех этих историях отчетливо проступают нерешенные задачи универсального характера.  
  
Аадхаар, или Кошмар с безопасностью  
  
Во всем, что касается озабоченности рядовых граждан своей приватностью как фундаментальным правом на тайну личной жизни, Россия гораздо больше похожа на Индию, нежели на страны Западной Европы или США. Это, как говорится, самоочевидный факт. Убедиться в котором обычно проще простого – достаточно лишь заглянуть в национальные СМИ и обратить внимание, насколько редко там обсуждаются проблемы вторжения государства и корпораций в область приватной жизни граждан.  
  
В Индии, однако, эта ситуация довольно резко переменилась в феврале нынешнего года – когда об откровенных посягательствах на приватность начали вдруг говорить и писать чуть ли не все. Не то чтобы причиной, а скорее поводом для давно назревшего обсуждения проблемы – сначала в обществе, а затем и в Верховном суде – стал один "твит".  
  
Суть данного твита, опубликованного 3 февраля 2017 года, заключалась в наглядной демонстрации того, насколько легко и быстро можно устанавливать ныне личности случайных людей из толпы. Это делается благодаря системе «Аадхаар», охватившей уже практически все население Индии централизованной базой данных, в которой накоплена биометрическая информация на более чем миллиард граждан страны.  
  
Название Aadhaar переводится как «фундамент» или «основы». А означает это то, что власти Индии в качестве основ своей государственности решили провозгласить создание общенациональной базы данных для биометрической верификации личности. Единой и тотальной базы на все население страны с гигантской численностью, 1,2 миллиарда человек, – начиная с младенцев и заканчивая глубокими стариками.  
  
Биометрическая основа Aadhaar задумана и реализована так, чтобы на каждого человека в базе имелись отпечатки всех пальцев на руках, снимок радужки глаза, цифровая фотография лица плюс текстовое описание особенностей в наружности резидента. Итогом же занесения людей в базу является выдача им уникального идентификационного номера на специальной Aadhaar-карте.  
  
Формально участие граждан в наполнении базы «Аадхаар» объявлено делом как бы добровольным. Однако в реальности вся эта система выстроена таким образом, что совершить практически любую необходимую по жизни официальную процедуру человек теперь может лишь в том случае, если предъявит свой ИД-номер и часть тела для сканирования-верификации – как правило, радужку глаза или отпечаток пальца. Биометрическая проверка личности увязана, в частности, с любыми услугами и получением буквально всего — от талонов на еду, социальных субсидий и топлива для кухни до водительских прав, расчетного счета в банке, кредитов, страховки, пенсии и так далее.  
  
Гигантская же проблема с Aadhaar заключается вот в чем. С одной стороны, никто в обществе толком не представляет, как устроено в этой системе обеспечение защиты, целостности и верификации данных. А с другой, регулярно появляющиеся в СМИ публикации свидетельствуют, что средства безопасности работают в базе не просто плохо, а очень плохо. Для примера, в данных «Аадхаар» уже многократно выявлены заведомо неверные и фальшивые записи. Под эту категорию, в частности, подпадают такие случаи, когда собственные ИД-номера и карточки оформлены на собак и коров. В другом нашумевшем случае было обнаружено, что персональная карта заведена на Ханумана, одного из богов индуистского пантеона.  
  
Даже неспециалисты понимают, что систематическое появление ложных данных в такой системе – это признаки создания новых каналов для отмывания денег и прочих форм мошенничества. Причина же очевидных проколов, судя по всему, заключается в том, что база данных Аадхаар никогда по-настоящему не проходила подобающие процедуры аудита и верификации.  
  
Так, во всяком случае, считают многие специалисты по инфобезопасности, а также юристы, правозащитники и политики, пытавшиеся ознакомиться с работой системы поглубже. Но многочисленные призывы оппозиции, обращения правозащитников и требования профессионалов инфозащиты к независимому и открытому аудиту Aadhaar остаются в правительстве, по сути дела, без ответа.  
  
Именно по этой причине общественность Индии сочла чрезвычайно важным нынешнее решение Верховного суда о том, что персональная приватность является фундаментальным гражданским правом каждого человека. А чтобы стало понятнее, отчего признанию этой идеи так активно сопротивляется правительство и почему данный конфликт теснейшим образом связан с развертыванием системы «Аадхаар», имеет смысл прояснить такие моменты.  
  
Общепринятая в мире юридическая практика такова, что признание фундаментальных прав людей на тайну приватной жизни сопровождается принятием соответствующего государственного закона о приватности и защите персональных данных. Причем в подлинно демократических странах (к которым Индия, безусловно, себя относит) законы подобного рода обычно требуют, чтобы все граждане, на которых собираются их персональные данные, включая и биометрию, не только знали, кто и зачем это делает, но и могли бы со своей стороны проконтролировать аккуратность и актуальность накапливаемой о них в базах данных информации.  
  
Ну а нынешняя ситуация в Индии такова, что в государстве до сих пор нет никакого четкого законодательства, которое могло бы защищать приватность и права миллиарда с лишним человек, предоставивших в базу Aadhaar свои данные. Мало того, что система работает абсолютно непрозрачно, так у людей при этом нет ни инструментов, ни возможностей каким-то образом контролировать или менять свои данные в базе в тех ситуациях, когда они оказываются скомпрометированы. Ибо жульничеством с биометрией уже никого не удивишь, кражи личности с целью мошенничества – в подобных централизованных базах дело обычное, а изменить свое лицо, радужку глаза или отпечатки пальцев граждане, ясное дело, уже не могут.  
  
И дабы этих неприятностей не показалось мало, еще одним очень существенным фактором в разворачивании Aadhaar стала энергичная коммерциализация системы. Иначе говоря, вокруг своей тотальной базы данных, охватившей все население, власти совершенно официально создали и активно наращивают специальную коммерческую инфраструктуру India Stack. Ее цель – обеспечивать разнообразные формы сотрудничества между правительственными ведомствами и частными компаниями, которым предлагается опереться на богатейшие возможности Аадхаар по персональной идентификации всех людей, внесенных в базу.  
  
Демонстрацией именно этих замечательных возможностей и была, собственно, нашумевшая twitter-публикация в феврале, отметившая таким образом плодотворное сотрудничество между India Stack и коммерческой фирмой OnGrid. Таких компаний, подобно OnGrid уже подключившихся к освоению Aadhaar, ныне насчитывается свыше двух сотен. А это означает, что у граждан Индии сегодня нет не только возможностей проконтролировать состояние и защиту записей о своей личности в центральной базе данных, но и никакой информации о том, сколько еще десятков или сотен компаний купили у правительства доступ ко всем их персональным данным ради наращивания собственных прибылей.  
  
Самое же примечательное, что наиболее характерные особенности – настойчивое желание сохранять непрозрачность системы Больших данных в сочетании с мощным коммерческим фактором – свойственны не только Индии, но и всем прочим новостям о проблемах в данной области.  
  
Большая шведская утечка  
  
В последних числах июля 2017 серьезнейший ИТ-скандал в Швеции заставил главу правительства Стефана Лёвена отправить в отставку двух министров – министра внутренних дел Андерса Игемана и министра инфраструктуры (транспорта) Анну Йоханссон. Блок оппозиционных партий требовал также отставки министра обороны Петера Хультквиста, замешанного в скандале не меньше уволенных коллег, однако его было решено оставить на посту – «дабы не вызывать общий правительственный кризис» в государстве.  
  
Из весьма скупых и лаконичных комментариев премьер-министра Лёвена на экстренной пресс-конференции по поводу отставок стало известно, что произошедшее – «это очень серьезное дело, подвергшее риску Швецию и жизни шведских граждан». Из-за неправильной работы с национальной информационной системой в бесконтрольном доступе оказались личные данные практически всех граждан Швеции, включая военных, сотрудников спецслужб и полиции.  
  
Хотя публично известно о данной утечке стало лишь только летом 2017-го, на самом деле история эта началась довольно давно, еще весной 2015 года. Более того, если судить формально, то можно говорить, что по большому счету она все еще продолжается. Причину такого казуса удобнее пояснить чуть позже. А пока куда полезнее просто обратить внимание на то, насколько больше власти здесь были озабочены сохранением своих проколов в тайне, нежели исправлением очевидно неблагополучной ситуации с защитой данных.  
  
Когда же о произошедшем все-таки стало известно народу, а за собственные расследования взялись журналисты и инициативные энтузиасты от общественности, то восстановленная ими суть истории стала выглядеть вкратце так. Фактически компрометация безопасности правительственной инфосистемы произошла в марте 2015 года, когда государственное Транспортное агентство всю работу по поддержанию своих баз данных переложило по коммерческому контракту на корпорацию IBM и на её технические ресурсы (включая людской персонал), базирующиеся в республике Чехия.  
  
Примерно

таким же аутсорсингом и размещением своих данных во внешних облачных сервисах занимаются ныне чуть ли не все, особенно, если хочется сэкономить. Однако именно в данном конкретном случае критично важным результатом предпринятого шага стало вот что. Огромные массивы чувствительной к разглашению информации, включая полный национальный реестр Швеции по водительским лицензиям и сведения о гражданах с засекреченными по той или иной причине персональными данными, оказались доступны посторонним людям, в основном иностранцам, которые не проходили никакой проверки в службах безопасности Швеции на предмет доступа к гостайне.  
  
Что именно там было скомпрометировано, в каких масштабах и с какими именно базами – практически всё это является предметом домыслов и фантазий, не подтверждаемых какими-либо конкретными документами. Ведь правительство и поныне изо всех сил старается сократить до минимума любые утечки о конкретных фактах произошедшего. Но вот в тех ситуациях, когда непосредственными свидетелями явно некомпетентного обращения с секретными сведениями оказывались очень многие коммерческие клиенты Транспортного агентства, утаивать очевидное становилось уже совершенно невозможно. Особую известность, в частности, обрел такой недавний случай.  
  
С наступлением весны, в марте, Агентство подготовило свежую версию полного реестра транспортных средств Швеции и их владельцев – распространив его среди всех маркетинговых и бизнес-структур, купивших подписку на данную рассылку. Само по себе это вполне нормально, поскольку реестр транспортных средств является публично доступной информацией. Но то две вещи с этой рассылкой пошли совсем не так.  
  
Во-первых, в открытый реестр, распространяемый для всех желающих по коммерческой рассылке, оказались включены люди, чье настоящее имя было изменено по программе защиты свидетелей и прочим подобным государственным программам.  
  
А во-вторых, когда эта грубейшая ошибка была обнаружена, то Агентство не стало рассылать новую, подчищенную версию реестра с инструкцией уничтожить прежний вариант как ошибочный. Вместо этого все уже и без того скомпрометированные имена были собраны в отдельный перечень, и теперь уже данный поименный список разослали еще раз открытым письмом по электронной почте – с требованием, чтобы подписчики сами занялись уничтожением в реестре всех записей на перечисленных в списке людей.  
  
Следует подчеркнуть, что это всего лишь один из достоверно известных ляпов. Сколько же их было всего и насколько серьезными была раскрываемые при этом сведения – такой информации публика не имеет. Но зато хорошо известно, кого и когда назначили главным виновником столь серьезных и систематических нарушений в правилах обращения с данными конфиденциального и секретного характера.  
  
В официальном заявлении нового руководства Транспортного агентства сообщается, что их бывшая генеральная директриса Мария Агрен одобрила контракт с IBM на аутсорсинг, хотя он нарушал сразу несколько законов страны: Закон о национальной безопасности, Закон о защите личных данных и Закон о публичности и приватности. В качестве наказания за допущенные нарушения госпожа Агрен в начале 2017 года была уволена и оштрафована на 70 000 крон «за беспечность с секретной информацией» (в пересчете на более известную валюту эта сумма составляет около 8,5 тысяч долларов и равна половине ее ежемесячной директорской зарплаты).  
  
За скупыми словами этой официальной информации, однако, очевидно скрывается куда более мутная и масштабная история. Независимые расследования показали, что шведская контрразведка, военные и прочие информированные люди в правительстве знали обо всех нарушаемых здесь законах страны если и не заранее, то как минимум с начала 2016 года. Однако наказать Марию Агрен решили только в январе 2017-го, когда случился особо серьезный прокол, о котором стало известно премьер-министру Лёвену.  
  
Глава же правительства, судя по всему, счел происходящее делом чрезвычайно серьезным и требующим пресечения и наказания. Так что уже в марте 2017-го дело бывшего гендиректора Агентства было рассмотрено в суде, так же стремительно был вынесен и вердикт о её «виновности». Вот только наказание за это уголовное преступление оказалось в высшей степени странным и непонятным – административный штраф в размере половины месячной зарплаты.  
  
Непонятными эти события, впрочем, выглядят лишь для тех, кто слабо разбирается в тонкостях шведской судебно-правовой кухни. А вот для людей, знакомых с правилами и особенностями юридических процедур, суть произошедшего представляется достаточно ясной. Поскольку наказание административным штрафом уже не может быть пересмотрено в случае признания вины обвиняемой стороной (а именно так и произошло), то теперь публика никак не сможет увидеть никаких документов, собранных и рассматривавшихся судом по этому делу. Иначе говоря, фактически тут был ловко и профессионально выполнен известный трюк типа «концы в воду»...  
  
Казалось бы, понятно, что директор государственного агентства совершила преступление. Она подписала документ, санкционировавший «исключение из текущего законодательства», – а это нечто такое, чего делать никаких прав у нее не было. Директор правительственного учреждения не может по собственному выбору подчиняться или не подчиняться закону. Однако от адвокатов Агрен стало известно, что с её точки зрения данная ситуация выглядела существенно иначе.  
  
Линия защиты на слушаниях в суде сводилась к тому, что «отход от текущего законодательства», который она лично санкционировала, не только считался вполне приемлемым в нынешних обстоятельствах, но и не был чем-то новым – «раньше такого рода решения уже принимались, поскольку это часть процесса, который позволялось применять при необходимости». Иначе говоря, если бы публика имела возможность копнуть данную историю поглубже, то стало бы понятнее, отчего именно Марию Агрен назначили тут крайней.  
  
Все примерно то же самое было в других ведомствах и раньше, ей же просто сильно не повезло, похоже, с тем, что «облако» IBM именно в данном случае оказалось расположено в Восточной Европе. Кроме того, по словам Агрен, она всегда работала в гражданском учреждении, а с вступлением на высокий пост с ней никто не проводил никаких специальных инструктажей об аспектах безопасности этой работы. Так что она в подобных условиях просто старалась делать всё так же, как и остальные, только еще лучше. Почему же спецслужбы Швеции, давно знавшие о том, кто и где работает с базами Транспортного агентства, только в 2017 году решили озаботиться подбором соответствующих кадров с допуском к гостайне – и обещают их укомплектовать «где-то к концу осени»? На этот естественный вопрос внятного ответа у властей нет.  
  
Из лагеря шведских правозащитников, активно выступающих за более эффективные инструменты, обеспечивающие право граждан на приватность, по поводу всей данной истории вполне резонно прозвучала следующая идея. Уж если государство оказывается столь неуклюжим и некомпетентным в защите собственных тайн и персональных данных о своих секретных сотрудниках, то что тут вообще можно говорить о защите государством персональной информации обычных граждан? Вполне возможно, что нечто компетентное и содержательное на данный счет могла бы поведать еще одна важнейшая фигура шведского скандала – американская корпорация IBM. Однако там, по давно уже заведенным известным правилам, полностью дистанцировались от внутренних политических разборок в скандинавском государстве и вообще отказываются от каких-либо комментариев к данной истории...  
  
Парадоксальный бизнес Palantir  
  
Случилось так, что корпорация IBM некоторым существенным, но одновременно незаметным образом оказывается непосредственно замешана и в третьей из разбираемых здесь новостей. Но здесь (точнее, в США) центральной фигурой конфликта являются не системы IBM, а их главный конкурент в области датамайнинговых услуг для властей – корпорация Palantir Technologies.  
  
Ну а поскольку две уже выявленных ранее проблемы в теме «государство и Большие данные» – непрозрачность системы и активная коммерциализация информации – в специфическом бизнесе Palantir проступают особенно ярко и отчетливо, имеет смысл рассмотреть эту историю как своего рода эталон. Или как архетипический образец, наглядно демонстрирующий не только корневые причины серьезных проблем, но и то, почему решить их в нынешних условиях, по сути, невозможно – на принципиальном уровне.  
  
Как многим, вероятно, известно, термин «палантир» появился в нашем лексиконе из толкиеновской эпопеи «Властелин колец», где данным словом именуют магические каменные шары, позволяющие их владельцам знать и видеть всё, что происходит в мире Средиземья. Аналитические компьютерные системы Palantir делают для их владельцев по сути то же самое – накапливая, просеивая и сопоставляя гигантские массивы самых разных данных и выстраивая на этой основе визуально внятные картины событий со взаимосвязями между их участниками.  
  
Подробности о создателях и лидерах, о занятной истории восхождения Palantir Tech и не менее любопытных идеологических основах этой фирмы можно найти в материале «Бизнес на мифах». Ну а здесь следует рассказать о тех особенностях ее продукции, благодаря которым бизнес молодой и «как бы либертарианской» компании стал гиперуспешным благодаря государственным заказам – с одной стороны. А с другой — уже успел ярко продемонстрировать, что опора властей на системы Palantir – это, очевидно, путь «не туда», если для государства важнее интересы общества, а не прибыли корпорации. В сообщениях из текущей ленты новостей, естественно, никто столь сильных обобщений не делает. Однако неоспоримые факты тут говорят сами за себя. Надо лишь соответствующим образом на эти факты смотреть. Иначе говоря, далее будут рассмотрены не просто свежие новости о Palantir, но и то, что они означают в более общем контексте.  
  
Одна новость — про руководство NYPD, крупнейшего в США регионального Управления полиции Нью-Йоркского мегаполиса, которое сильно повздорило с корпорацией Palantir из-за ее «волшебной» системы анализа. Точнее говоря, причиной конфликта стала не собственно система, уже несколько лет применяемая в NYPD, а трудности перехода полицейских к более дешевому и более удобному для них альтернативному ПО, разработанному корпорацией IBM.  
  
Причина же затруднений здесь такова, что всю ту бездну информации, которая была наработана полицией в их системе за время эксплуатации, Palantir категорически отказывается переносить в систему IBM, ссылаясь на защиту своих ценнейших «коммерческих тайн». Точнее говоря, они с готовностью и сразу выдают все первичные или сырые данные, на основе которых база наполнялась. Но вот самое главное – собственно аналитику – выводить для переноса отказываются наотрез. Ибо это раскроет их методы обработки, то есть главную интеллектуальную собственность корпорации.  
  
Суть обозначившегося здесь конфликта имеет смысл обозначить как можно отчетливее. Формально имеется работающая исключительно на общество правоохранительная структура, а также есть некая бизнес-фирма, технически обслуживающая эту структуру в своих сугубо коммерческих интересах. В реальном же итоге оказывается так, что вся действительно необходимая для работы информация, формально собранная и как бы принадлежащая полиции, на самом деле не только принадлежит бизнес-фирме, но и отдавать её на защиту общества она решительно не желает – по причине своих сугубо шкурных коммерческих интересов...  
  
Новость другая – о проблемах с системой Palantir у полицейских LAPD, другого крупнейшего в США Управления полиции Лос-Анджелеса. Проблемы тут обозначились существенно иного рода, однако и они очень наглядно рисуют, сколь опасным сочетанием оказываются для государства непрозрачность и коммерция при анализе Больших данных.  
  
Одним из главных факторов для аналитических успехов системы Palantir является как можно более широкий охват разнообразных источников входной информации. В частности, для эффективной работы с полицейскими базами данных оказывается очень полезным просеивать и сопоставлять не только записи разных баз одного региона, но и одновременно выявлять их взаимосвязи с записями в базах других регионов. Поскольку при таком расширении общей базы не только повышается эффективность полицейской аналитики, но и заметно растут прибыли корпорации, в Palantir очень энергично работают над расширением клиентов в регионах, предоставляя существенные скидки всем новичкам.  
  
Столь замечательная бизнес-стратегия, однако, находится в сильнейшем противоречии с известным принципом «секционирования» информации, общепринятым в работе структур, имеющих дело с секретными сведениями. Проще всего суть этого принципа можно пояснить на основе прямоугольной таблицы, где строки соответствуют разным уровням допуска сотрудников к гостайне – «для служебного пользования», «секретно», «совершенно секретно», «особой важности». А колонки таблицы – это разные секретные дела или «объекты», с которыми работает именно данное ведомство. Тогда любой из сотрудников этого ведомства имеет доступ лишь к той информации, которая находится в клетках или «секциях» таблицы, соответствующих (а) уровню их общего доступа к гостайне и одновременно (б) разрешению на доступ к сведениям конкретно по данному объекту.  
  
Официально провозглашается, что в аналитических программах Palantir эти ясные принципы реализованы совершенно замечательно. Здесь не только имеется строгая система допусков, гарантирующая, что всякий сотрудник может получить доступ к данным исключительно на уровне своих полномочий, но и автоматически создается аудиторский след, в котором фиксируется всё – кто именно видел определенные фрагменты информации и что именно с этими фрагментами делалось.  
  
В реальной работе лос-анджелесской полиции, однако, вся эта картина выглядит существенно иначе (наверняка то же самое происходит и у других пользователей системы, но их недовольство просто не доходит до прессы). Сотрудники спецподразделения LAPD, занимающиеся борьбой с организованной преступностью, на протяжении многих и многих месяцев безуспешно пытались добиться от Palantir всего лишь того, чтобы особо чувствительные файлы о людях и операциях, которые они помечают конфиденциальными, не становились свободно доступными для всех остальных пользователей системы. Однако даже эта тривиальная, казалось бы, техническая проблема, оказывается практически неразрешимой, когда для корпорации-подрядчика гораздо важнее максимально широкий охват анализируемой информации, нежели узкие интересы копов-оперативников...  
  
Описанные здесь проблемы – это не что-то новое и неожиданное. Все подобные трудности с большими базами данных известны профессионалам очень давно. Как они выражаются, «приватность масштабируется очень плохо». И если в работе секретных спецслужб об этой серьезнейшей беде широко вещать обычно не принято (как и обо всех прочих гостайнах), то в работе всеобщей системы здравоохранения – где в обращении с деликатными данными пациентов прослеживается много аналогий – о том же самом специалисты по инфозащите говорят давно и громко.  
  
Медицинские базы данных с цифровыми картами пациентов очень полезны врачам, и все там работает замечательно – когда штат из 20 человек в медицинском учреждении имеет доступ к медкартам десяти тысяч пациентов, зарегистрированных в их картотеке. Но вот когда создается единая централизованная система, позволяющая каждому доктору или медсестре данной страны просматривать медицинские записи любого пациента – вот это всегда получается однозначно плохо, организовано неправильно и тут же ведет к злоупотреблениям разной тяжести (от сливов болезней знаменитостей в таблоиды до отказов людям в работе, кредитах или в страховке на основании данных об их здоровье в цифровых медкартах, считающихся как бы «приватной информацией», а на самом деле доступных кому угодно.

Всем понятно, что решать проблему по-любому необходимо – все базы Больших данных постоянно только прирастают, а отказываться от их преимуществ совершенно никак невозможно. Но одновременно понятно и то, что простых и эффективных решений c защитой здесь, похоже, не существует. Ярчайшим подтверждением тому являются регулярные и мощные сливы информации из наиболее секретных недр разведслужб – начиная от демарша Эдварда Сноудена и заканчивая массами свежих публикаций о хакерском инструментарии АНБ и ЦРУ США. Причем сами спецслужбы оказываются зачастую не в силах как определить то, кто именно стоит за компрометацией, так и оценить объемы скомпрометированных данных.  
  
Отсюда понятно, наверное, что с защитой приватности для всех обычных граждан государство разберется не раньше, чем решит проблемы с массовыми утечками компромата из спецслужб. А это две стороны одной и той же проблемы...  
  
Бёрд Киви

# Big Data уже нельзя назвать инновацией

**18 ноября 2015**

Директор по маркетингу сервисов «Яндекса» и преподаватель курса «Big Data для руководителей» Школы новых медиа НИУ ВШЭ рассказал Cossa о том, что происходит сейчас в области больших данных.

— **Некоторые эксперты**[**утверждают**](http://www.infoworld.com/article/2907872/big-data/big-data-is-dead-long-live-big-data.html)**, что эра Big Data подходит к концу. Вы с этим согласны?**

— Нет, не согласен. Если внимательнее посмотреть на мнение уважаемых экспертов, например, из компании Gartner, то окажется, что к концу подошло совсем другое: эпоха пребывания Big Data в категории «развивающиеся и многообещающие технологии». Согласно оценкам агентства, в 2015 году использование Big Data стало частью базовых технологий в более чем 20% компаний. Большие данные стали настолько широко распространенными, что их уже нельзя назвать инновацией. Впрочем, эта статистика не относится к России.

— **Есть ли разница между Big Data и Data Science?**

— Мне представляется, что есть. В словах Big Data больше маркетинга, чем содержания: это словосочетание позволяет продавать громоздкие и дорогие решения. Data Science — это современный способ использовать разнообразные данные в бизнесе, выходящий за пределы привычной статистики или бизнес-аналитики. Применять подходы Data Science можно и в тех случаях, когда размер доступных данных не является фантастическим.

— [**Есть мнение**](http://www.infoworld.com/article/2907872/big-data/big-data-is-dead-long-live-big-data.html)**, что следующим скачком будет Data Driven.**

— Смысл использования Data Science как раз в том, чтобы данные становились драйвером решений в бизнесе, драйвером изменений в бизнес-процессах. В этом смысле компания, в которой практикуется реальное использование Data Science, автоматически становится Data Driven.

—  **Business Insider**[**писал**](https://www.google.com/url?q=http://www.businessinsider.com/technologies-worth-billions-in-2015-2014-12&sa=D&ust=1447404920378000&usg=AFQjCNEuLniMdGt6qE8FF8PBtp6U7kqWcg)**, что от 2015 стоило ожидать новых способов анализа Big Data, в частности, появления инструментов для анализа визуальных источников информации. Существуют ли уже на рынке подобные технологии?**

— Инструментов анализа визуальной информации существует достаточно много. Важно не столько их присутствие на рынке, сколько коммодитизация: все большее число необходимых программ становятся доступными в открытом доступе, потребные вычислительные мощности можно арендовать. Это означает, что реальное использование Data Science становится доступно не только крупным компаниям с большими собственными ресурсами, человеческими и машинными, но всем нуждающимся.

— **Есть ли на рынке квалифицированные специалисты в области Big Data?**

— Есть, но дефицит очень силен. Причем в дефиците не только технические специалисты, способные организовать саму работу с данными, но и менеджеры, понимающие как интегрировать Data Science в бизнес-процессы, как правильно нанять специалистов, как правильно ставить им задачи и организовывать их работу. Собственно, последних мы и готовим в Школе новых медиа, так как именно их особенно не хватает на российском рынке.

— **Какие люди требуются в этой области?**

— Область достаточно широка, и в ней требуются разные специалисты, но всех их объединяет хорошая математическая подготовка и достаточно высокий уровень владения языками программирования. В зависимости от специфики конкретных задач могут потребоваться дополнительные знания и умения: хороший data scientist — довольно синтетический специалист.

— **В какую сторону, на ваш взгляд, будет развиваться рынок Big Data в дальнейшем?**

— В сторону вытеснения рутинного умственного труда трудом машин. Причем в первую очередь в промышленности, а не только в быту.

[**Вернуться в каталог сборников по менеджменту**](http://учебники.информ2000.рф/management2/management3.shtml)

[**Вернуться в электронную библиотеку по экономике, праву и экологии**](http://учебники.информ2000.рф/index.shtml)

[**НАПИСАНИЕ на ЗАКАЗ: дипломы и диссертации, курсовые и рефераты. Переводы с языков, он-лайн-консультации. Все отрасли знаний**](http://учебники.информ2000.рф/napisat-diplom.shtml)

|  |  |
| --- | --- |
| [**КНИЖНЫЙ МАГАЗИН**](http://учебники.информ2000.рф/chitai.shtml) |  |

|  |  |
| --- | --- |
| [**ТОВАРЫ для ХУДОЖНИКОВ и ДИЗАЙНЕРОВ**](http://учебники.информ2000.рф/kar.shtml) |  |

|  |  |
| --- | --- |
| [**АУДИОЛЕКЦИИ**](http://учебники.информ2000.рф/lectr.shtml) |  |

|  |  |
| --- | --- |
| [**IT-специалисты: ПОВЫШЕНИЕ КВАЛИФИКАЦИИ**](http://учебники.информ2000.рф/otu.shtml) |  |

|  |  |
| --- | --- |
| [**ФИТНЕС на ДОМУ**](http://учебники.информ2000.рф/fit1.shtml) |  |