

Большие данные в развитии платформы

Самойлов Никита Константинович
Студент 4-го курса факультета управления Санкт-Петербургского
государственного экономического университета;
e-mail: konstantintinovich@gmail.com

Шкарлат Ольга Сергеевна
Студентка 4-го курса факультета управления Санкт-Петербургского
государственного экономического университета;
e-mail: olia.shkarlat@yandex.ru

Степаненко Дарья Александровна
К.э.н., доцент кафедры менеджмента и инноваций
Санкт-Петербургского государственного экономического университета;
e-mail: kadar_77@mail.ru

Аннотация. В исследовании рассматриваются облачные платформы, применение Big Data в современных компаниях и влияние больших данных на развитие цифровых платформ. Даны понятия и сравнительная характеристика наиболее часто встречающихся облачных платформ – частных и публичных. В статье определены условия эффективного использования платформенных типов ведения бизнеса. Изучен опыт использования Big Data в международной и отечественной компаниях. На основе выявленных задач Big Data сделаны выводы о предпочтительном применении частных облачных платформ крупными предприятиями и использовании публичных облачных платформ малыми и средними предприятиями. Исследовано положительное и отрицательное влияние использования платформенных моделей ведения бизнеса и анализа больших данных. Сделаны предположения о перспективах развития частных и публичных облачных платформ в России и мире.

Ключевые слова: Big Data, большие данные, цифровые платформы, публичная облачная платформа, частная облачная платформа, цифровизация, облачные сервисы.

BIG DATA IN THE DEVELOPMENT OF THE PLATFORM

Stepanenko Daria A.
Senior lecturer, the Department of Management and Innovation,
St. Petersburg State University of Economics, PhD in Economics, St. Petersburg;
e-mail: kadar_77@mail.ru

Samoilov Nikita Konstantinovich
Student 4th year, Management department,
St. Petersburg State University of Economics, St. Petersburg;
e-mail: konstantintinovich@gmail.com

Shkarlat Olga Sergeevna
Student 4th year, Management department,
St. Petersburg State University of Economics, St. Petersburg;
e-mail: olia.shkarlat@yandex.ru

Abstract. The study examines cloud platforms, the use of Big Data in modern companies and the impact of big data on the development of digital platforms. We give concepts and comparative characteristics of the most common cloud platforms – private and public. The article defines the conditions for the effective use of platform types of doing business. We have been studied the experience of using Big Data in international and domestic companies. Based on the identified Big Data tasks, conclusions were drawn about the preferred use of private cloud platforms by large enterprises and the use of public cloud platforms by small and medium-sized enterprises. The positive and negative impact of the use of platform business models and big data analysis has been investigated. We made assumptions about the prospects for the development of private and public cloud platforms in Russia and the world.

Keywords: Big Data, big data, digital platforms, public cloud platform, private cloud platform, digitalization, cloud services.

Диджитализация коснулась всех сфер жизни общества, включая экономику, политику, социальную среду и пр. Это обусловлено массовым характером компьютеризации различных видов деятельности и отдельных технологических цепочек. В настоящее время можно выделить три направления цифровизации:

- в бизнес-среде;
- в органах государственной власти;
- в ожиданиях потребительских групп, исходя из понимания качества и количества необходимых благ [3].

Облачные платформы пользуются популярностью ввиду скорости создания новых приложений, внедрения эффективных инструментов анализа Big Data, гибкости и масштабируемости системы, машинного обучения и искусственного интеллекта. Синтез методов, инструментов анализа больших данных и использование облачных платформ в настоящее время помогает лидерам IT-компаний и не только строить ближайшие к достоверному прогнозы, делать выводы, приводящие компании к значительному росту прибыли, росту доли рынка и увеличению конкурентоспособности фирмы.

Большие данные (big data) – серия подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объемов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста.

Описанные в законе Мура возможности вычислительной техники, делают критерий «объем данных» неточным. Например, переход от измерения больших данных в терабайтах к петабайтам уводит их от категории «большие». Следовательно, главными критериями оценки больших данных должны стать их структурированность и вариативность представления.

Для анализа Big Data, информации, например, непрерывно поступающей с видео-, аудиорегистраторов и датчиков, мессенджеров и соцсетей, данных метеорологических служб, геолокации абонентов сотовой связи и пр., необходимо ее соответствие пяти главным характеристикам (рис. 1).

Рис.1. Главные характеристики Big Data



Таким образом, источниками Big Data могут служить:

- данные приборов учета (показания регистраторов, датчиков, сенсоров и пр.);
- корпоративные источники (архивные данные, информация баз данных и файловых хранилищ, транзакции и пр.);
- данные Internet (интернет вещей, информация соцсетей, блогов, форумов и пр.) [2].

Многообразие платформ отражено в организационных возможностях компаний:

- создание собственной платформы, используя физическую инфраструктуру;
- аренда мощностей провайдера;
- совмещение собственной платформы и арендованных мощностей.

Для аналитики больших данных в мировой практике наиболее часто используют частные и публичные облачные платформы. Для частной облачной платформы характерно:

- локальное расположение дата-центра компании или у стороннего поставщика;
- обязательное размещение инфраструктуры в частной сети;
- программная и аппаратная обеспеченность в рамках одной компании.

Действуя на основе Федерального закона «Об информации, информационных технологиях и о защите информации» от 27.07.2006 № 149-ФЗ [1], данный тип платформ развивают крупные организации – органы государственной власти, медицинские и финансовые организации, обязанные хранить данные у себя.

Публичная облачная платформа – это готовая платформа для анализа Big Data, принадлежащая провайдеру и предоставляющая облачную инфраструктуру стартапам и развивающимся компаниям (облачная платформа как услуга). Таким образом одновременно могут пользоваться несколько или много компаний, но каждая компания может получить доступ только к своим данным.

При выборе частной или публичной облачной платформы для работы с большими данными применяют следующие характеристики эффективности [5]:

Экономичность (финансы). Для развертывания частной облачной платформы необходимо приобрести оборудование, нанять персонал, развернуть и обслуживать инфраструктуру – это очень большие расходы. При публичной облачной платформе никаких первоначальных вложений нет. Первоначальные вложения отсутствуют, необходима только плата за мощности, которые компания использует.

Масштабируемость. В *частной* облачной платформе возможности масштабирования развернуты на оборудовании и ограничены самим же оборудованием. Публичная облачная платформа может подстраиваться под изменения, в связи с чем возможно выделение больших мощностей для обработки и хранения данных. При отсутствии ресурсов для анализа больших данных мощности облачной инфраструктуры не используются.

Экономичность (время). *Частную* облачную платформу компаниям нужно обслуживать и администрировать самостоятельно. В *публичной* облачной платформе провайдер занимается обслуживанием системы обработки данных, что позволяет компании повысить эффективность аналитики.

Быстрый запуск проекта (time-to-market). *Частная* облачная платформа требует больших инфраструктурных мощностей и капитальных затрат, из-за чего может замедлиться выпуск IT-продуктов на рынок. *Публичная* облачная платформа позволяет ввести в эксплуатацию IT-инфраструктуру без крупных первоначальных вложений. В общедоступном облаке инфраструктура аналитики данных настраивается за часы.

Отказоустойчивость. В *частной* облачной платформе бесперебойная работа требует больших капитальных вложений, так как она обеспечивается средствами Disaster Recovery. В *публичной* облачной платформе провайдер использует средства для обеспечения бесперебойной работы, значительно снижающие время сбоев системы.

Обеспечение требований законодательства. В *частной* облачной платформе компания следит за выполнением требований законодательства и регуляторов. В *публичной* облачной платформе ответственность за соблюдение законодательства, требований и стандартов лежит на провайдере.

Примерами эффективного использования Big Data являются отечественные и зарубежные компании, среди которых российская сеть гипермаркетов HOFF и авиакомпания S7 Airlines, ПАО «Газпромнефть», американская транснациональная компания PepsiCo и британско-нидерландская нефтегазовая компания Royal Dutch Shell.

Сеть гипермаркетов HOFF – сеть магазинов мебели и аксессуаров для дома, функционирующих в формате гипермаркета в России. Гипермаркет HOFF работает по принципу «Все в одном месте». Бизнес-задачи компании нацелены на внутренний рост конверсии сайта и количества мультиканальных покупателей.

Для решения поставленных задач все данные компании были аккумулированы в одном месте, разработаны собственные параметры эффективности, сформированы отчеты, управление ставками осуществлялось на основе отчетов компанией Alytics.

Результаты проведенных мероприятий:

- рост показателя ROI до 17% по некоторым позициям товаров,
- рассылка по электронной почте показала, что на 1 руб. онлайн-выручки приходится 4 руб. в офлайне,
- выросла значимость мультиканальных потребителей,
- в гипермаркетах Москвы треть прибыли принесли пользователи сайта Hoff.ru.

В российской авиакомпании S7 Airlines, второй крупнейшей в стране, выполняющей внутренние и международные пассажирские перевозки, входящей в топ-10 самых пунктуальных европейских авиакомпаний, бизнес-задачи сформулированы как «увеличение конверсий, снижение стоимости расходов на конверсии».

Для решения поставленных задач были созданы персонализированные креативы на основе Big Data. Повысилась точность прогнозов за счет накопления данных. Результатом стал рост конверсии в два раза при аналогичном бюджете и снижении стоимости конверсии на 40%.

В PepsiCo Big Data использовали для привлечения приоритетных покупателей, что обеспечило увеличение объема реализации на 79,8% в первые 3 месяца после запуска решения.

Анализ больших потоковых данных в крупнейшей нефтегазовой компании Shell использован для предвидения сбоев в работе оборудования и своевременного устранения неполадок, сокращения времени на аналитику складских запасов с двух дней до 45 минут. Сокращение расходов на перемещение и перераспределение ресурсов позволило компании сэкономить миллионы долларов в год.

Аналогичные результаты были получены в ПАО «Газпромнефть», где анализ более двухсот миллионов альтернативных показаний на 1649 скважинах в год позволил выявить причины сбоев насосов и обратного слива нефти при отключении насосного оборудования.

При анализе Big Data могут возникать ошибки. К типичным проблемам можно отнести [6]:

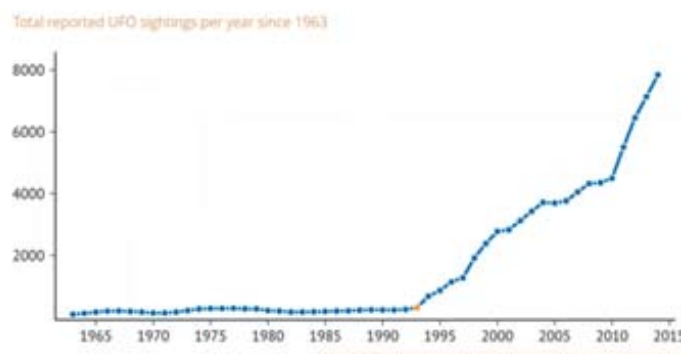
- ошибки выборки;
- ошибки корреляций;
- потеря данных;
- ошибки анализа.

Примерами неудачного использования больших данных с ошибкой выборки является опыт компании Street View, которая предложила следить за состоянием дорог жителям Бостона через мобильное приложение. Программа анализировала отклонения положения мобильного телефона от нормы, т.е. учитывала ямы, выбоины и др. дефекты дорог и отправляла данные в городские службы.

Анализ жалоб показал, что большинство из них получено от жителей престижных районов, а не бедных, именно жители богатых районов являлись активными пользователями приложений при постоянном подключении к Internet. В исследовании значимой единицей оказалось событие в приложении, а не человек, обнаруживающий изъяны на дорогах. Результатом анализа больших данных стало искаженное представление о состоянии дорог из-за демографии пользователей смартфонов.

Примером ошибки корреляций служит ситуация с резким увеличением количества наблюдений НЛО. На рис. 2 показано общее количество наблюдений НЛО с 1963 по 2015 г.

Рис. 2. Динамика общего количества наблюдений НЛО, 1963–2015 гг.



На диаграмме видно, что с 1993 г. инопланетяне начали активное изучение Земли, но понимание факта выхода первого эпизода фильма «Секретные материалы» дает объяснение такого экспоненциального роста.

Значительной ошибкой при анализе данных является их потеря. Например, профессор С. Левитт, один из авторов «Фрикономики», признал, что вывод о падении подростковой преступности в США из-за легализации абортов, а не из-за роста экономики и культуры был ошибочным, т.к. часть собранных данных была утеряна.

Кроме ошибок в обработке и анализе данных, следует учитывать и недостатки использования цифровых платформ.

1. Платформенная модель ведения бизнеса является разновидностью аутсорсинга, следовательно, возникает опасность передачи данных сторонней организации.
2. Угроза разглашения коммерческих данных относится к конфиденциальной информации, из-за чего возникают сложности с быстрым запуском и развитием компании в облачной платформе.
3. Создание облачных программ, являющихся важной частью бизнеса, предполагает работу компании на облачной платформе. При переходе на другую платформу или внутриорганизационную базу данных у владельцев торговых марок возникают дополнительные затраты на программное обеспечение. Затраты на переход с облачных технологий на внутриорганизационную базу данных неизбежны, но следует минимизировать влияние данного фактора.

Таким образом, можно выделить ряд задач, которые решает Big Data:

- составление более детализированного портрета целевой аудитории и клиентов, аккумуляция данных о причинах ухода потребителей и сегментирование клиентов;
- персонализация предложений, целевая ориентация рекламы;
- сбор информации о пользе продукта и его безопасности;
- оценка качества услуг, оборудования и улучшение логистики;
- упорядочение и удешевление хранения данных по заданным параметрам;

При анализе деятельности облачных платформ были выделены следующие положительные и отрицательные стороны (табл. 1).

Таблица 1. Достоинства и недостатки анализа больших данных.

+	–
Экономичность	Ошибки анализа и обработки данных
Масштабируемость	Недостатки аутсорсинга
Эффективность	Угроза разглашения конфиденциальной информации при быстром запуске бизнеса
Безопасность	Зависимость от облачной программы

Несмотря на наличие как положительных, так и отрицательных сторон анализа больших данных в рамках частных и публичных облачных платформ, платформенная модель ведения бизнеса будет стремительно развиваться. Существуют предпосылки существенных различий в развитии отечественных и зарубежных платформ, обусловленные готовностью к переходу на цифровые платформы. Лидерство в данной области сохраниться за США, Китаем, Великобританией, Индией и Германией [7].

Развитию частных и публичных цифровых платформ в России будет способствовать уточнение международной и национальной законодательной базы в области платформенной экономики со стороны государства и повышение цифровой грамотности и приобретения цифровых компетенций представителями бизнеса и физическими лицами.

Список литературы

1. Федеральный закон «Об информации, информационных технологиях и о защите информации» от 27.07.2006 № 149-ФЗ. Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_61798/
2. Вичугова А. Большие данные. Режим доступа: <https://www.bigdataschool.ru/wiki/%D0%B1%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D0%B5-%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5-big-data> (дата обращения: 12.05.2021 г.)
3. Салимьянова И.Г., Зинчик Н.С., Погорельцев А.С. и др. Инновационное развитие предприятий в условиях цифровой трансформации экономики: монография под общей редакцией д-ра экон. наук, проф. А.Г. Бездудной. – СПб.: Изд-во СПбГЭУ, 2020. 166 с.
4. Чуркина Н.С., Степаненко Д.А. Развитие сетевых структур: от внутренних рынков к платформизации // Стратегии бизнеса. Т.8, №8 (2020). С. 219–222 DOI: 10.17747/2311–7184–2020–8–219–222
5. Кушнир Е. Анализ больших данных в облаке: как бизнесу стать дата-ориентированным // Mail.ru Cloud Solutions. URL: <https://mcs.mail.ru/blog/analiz-bolshih-dannyh-v-oblake> (дата обращения: 13.05.2021 г.)
6. Большие ошибки в больших данных: проблемы анализа на практике. URL: <https://habr.com/ru/company/mailru/blog/512714/> (дата обращения: 13.05.2021 г.)
7. Доклад компании Accenture «Five Ways to Win with Digital Platforms» (2018) // URL: https://www.accenture.com/us-en/_acnmedia/PDF-29/Accenture-Five-Ways-To-Win-With-Digital-Platforms-Full-Report.PDF (дата обращения: 11.05.2021)