

# Формализация информации и Big Data

«02.03.03 - Математическое обеспечение и администрирование информационных систем  
направленность разработка и администрирование информационных систем»

<http://vikchas.ru>

Тема 2. Big Date

Лекция 12 «Большие данные и перспективы их развития»

Часовских Виктор Петрович

д-р техн. наук, профессор кафедры ШИиКМ

ФГБОУ ВО «Уральский государственный экономический  
университет»

Екатеринбург 2022

# Подходы к монетизации данных

<p>Кому принадлежат данные</p> <p>Как монетизировать</p>	 <p><b>ВНУТРЕННИЕ</b> данные компании</p>	 <p><b>ЧУЖИЕ</b> данные (внешние)</p>
<p><b>ТРАТИТЬ МЕНЬШЕ</b></p> 	<p><b>Эффективность производства</b></p> <p>Повышение эффективности бизнес-процессов за счет собственных данных</p>	<p><b>Эффективность производства. Коллаборация рыночных поставщиков</b></p> <p>Использование чужих данных, чтобы эффективнее организовывать процессы внутри компании</p>
<p><b>ЗАРАБАТЫВАТЬ БОЛЬШЕ</b></p> 	<p><b>Next best offer, cross-sales</b></p> <p>Больше продавать за счет того, что клиенту делают лучшее предложение, основанное на данных компании; разработка нового, более ценного продукта, таргетинг</p>	<p><b>Интернет-маркетинг, телемаркетинг</b></p> <p>Покупка аналитических сервисов, данные маркетинговых площадок</p>
<p><b>НОВЫЙ БИЗНЕС</b></p> 	<p><b>Аналитика, data-сервисы</b></p> <p>На основании собственных данных открывается новый бизнес или новые data-сервисы</p>	<p><b>Компетенции, сервисы, ИТ</b></p> <p>Монетизация компетенций, инфраструктуры</p>

# На рынке больших данных существуют четыре основные бизнес-модели

## Бизнес-модель

## Описание

## Аналоги на других рынках



**Подписка**

С покупателя ежемесячно или ежегодно взимается фиксированная сумма, независимо от частоты использования продукта

**ökko**

A



**Доля прибыли**

Покупатель выплачивает согласованную долю увеличения прибыли, достигнутого за счет внедрения решения, в течение определенного периода / существования совместного предприятия с согласованной схемой выплат

**Яндекс Такси**



**Плата за объем**

С покупателя постоянно и регулярно взимается плата за использование продукта (например, плата за мобильный интернет за Мб и т. п.)

**МЕГАФОН**



**Единовременная плата**

Покупатель платит только один раз за постоянный доступ к продукту

**1e®**

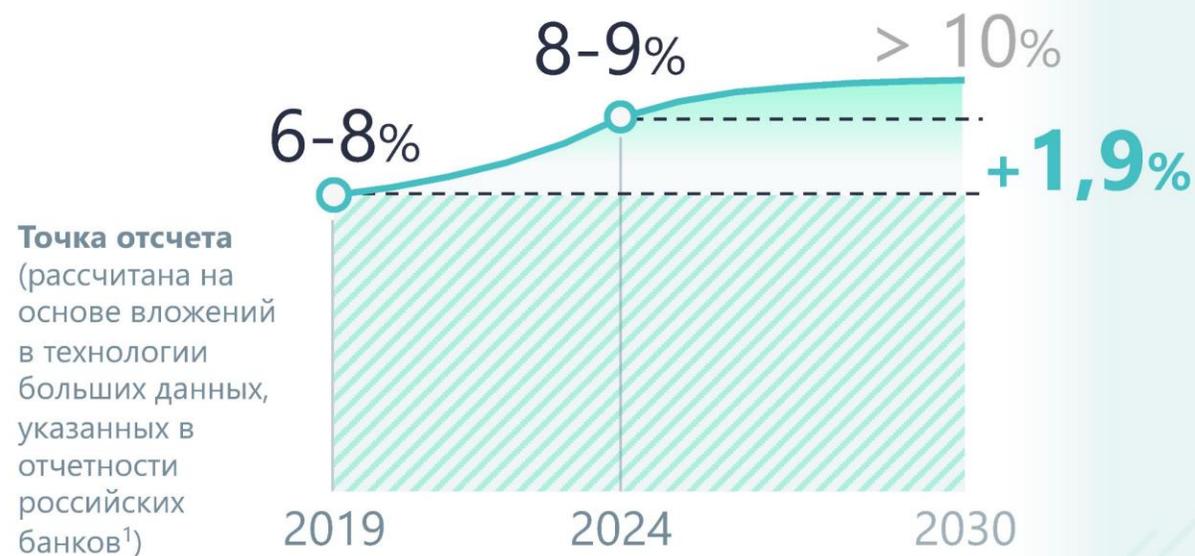
# Монетизация данных в финансовых институтах может дать дополнительный рост отрасли на 1,9% к 2024 году



**Источник:** рыночная стратегия по большим данным до 2024 года, разработанная Ассоциацией больших данных



## Прирост добавленной стоимости (%) от внедрения больших данных



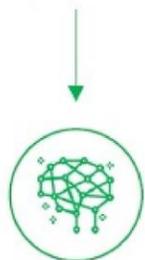
**Источник:** Министерство экономического развития, опыт АБД по работе над кейсами, разговоры с экспертами, анализ открытых ресурсов, выполненный АБД  
1. Включая финансовые отчеты и новостные статьи

# Монетизация данных. Кейсы

## Определение «справедливой ставки аренды» на базе модели машинного обучения

### ДАННЫЕ:

- Информация об объектах аренды и продажи (более 200 источников, в т.ч. региональных)
- Стоимость аренды залоговой недвижимости Сбербанка
- Стоимость аренды отделений Сбербанка



### MACHINE LEARNING

Определение аномалий в ставках аренды недвижимости и подбор наиболее перспективных объектов локации с помощью машинного алгоритма

The screenshot displays a real estate application interface. On the left, there are filter panels for 'Рыночные объекты аренды', 'Рыночные объекты продажи', 'Сбербанк - аренда', and 'Сбербанк - собственность'. The main area shows a map with various colored pins. A pop-up window for 'Рождественская, 33' is open, showing details for 'Сбербанк' and 'Рынок' properties. The 'Сбербанк' section includes a table with the following data:

Номер	Сегмент	Тип права	S, m <sup>2</sup>	Цена в мес.	Сравнение	Аналоги
1101/21	Офис	Аренда	67 m <sup>2</sup>	110 000 Р		
1101/18	Паркинг	Аренда	67 m <sup>2</sup>	110 000 Р		

The 'Рынок' section includes a table with the following data:

Номер	Сегмент	Тип права	S, m <sup>2</sup>	Цена в мес.	Сравнение	Аналоги
1101/21	Магазин	Аренда	66 m <sup>2</sup>	115 000 Р		
1101/18	Паркинг	Аренда	67 m <sup>2</sup>	110 000 Р		

At the bottom right, there is a 'Фильтры аналогов' section with filters for 'Аренда' and 'Продана', and a table of similar properties with columns for 'Entity type', 'Main type', and 'Локация'. The bottom status bar shows 'Количество объектов: 192', 'В расчете: 191', and 'Ср. стоимость: 1 300 000 Р'.

# Монетизация данных. Кейсы

## Прогноз выручки и потенциал локации для открытия торговой точки

Оценка экономической привлекательности отдельной локации на основе данных



### ЧТО АНАЛИЗИРУЕМ:

- Профиль потребителей и их доходы
- Покупательную способность
- Транзакционную активность
- Конкурентное окружение
- Стоимость аренды



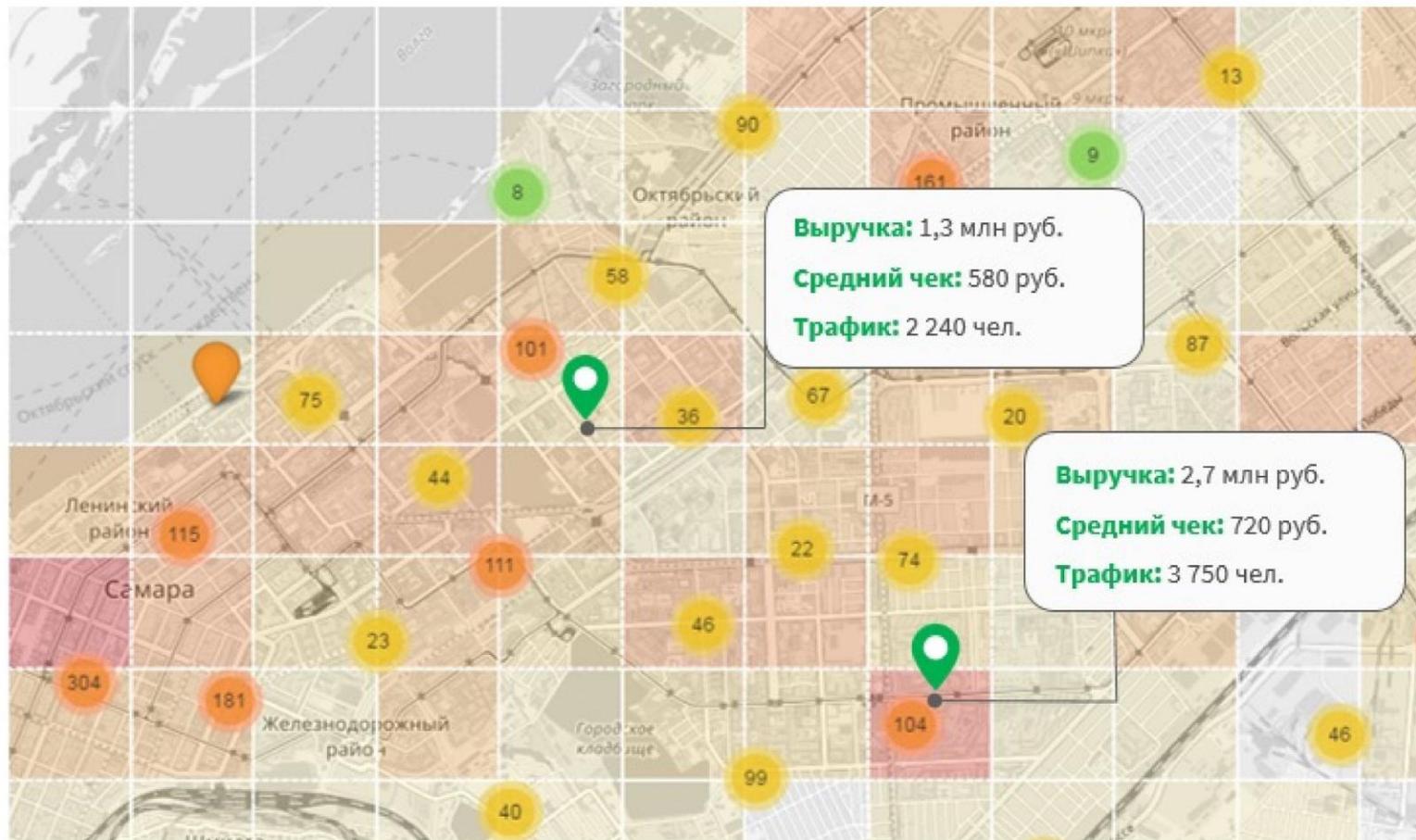
### ДЛЯ ЗАДАННОЙ ЛОКАЦИИ МЫ ПРОГНОЗИРУЕМ:

- Выручку торговой точки
- Трафик
- Средний чек



### ДАННЫЕ ПРИМЕНЯЮТСЯ:

- В бизнес планировании
- Для развития сети
- Для принятия решений об инвестициях (инвестиционные комитеты)

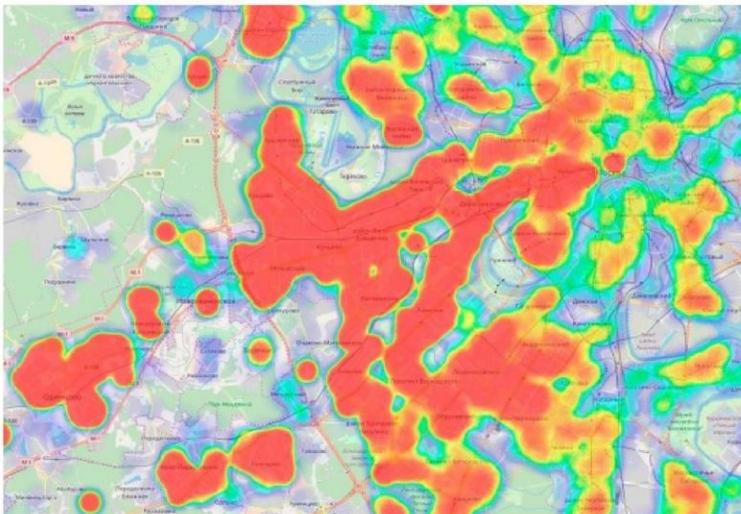


# Монетизация данных. Кейсы

## Аналитика торговых центров для повышения трафика покупателей

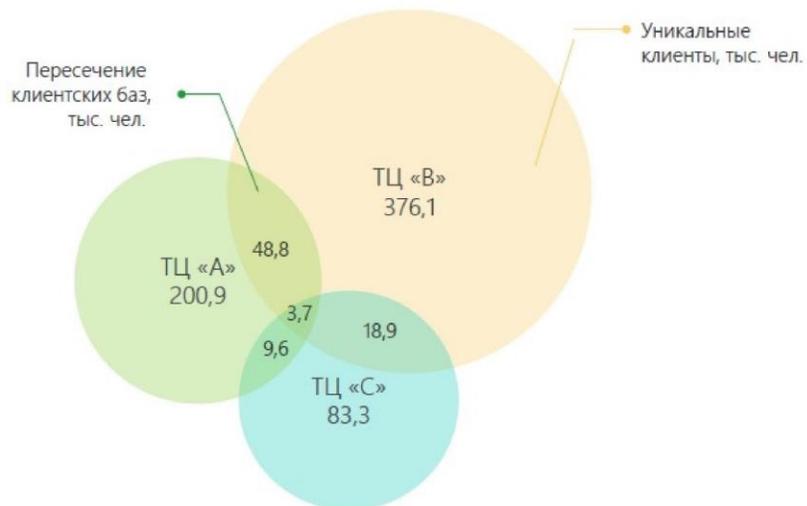
### Зона охвата ТЦ

Число клиентов в зоне шаговой и транспортной доступности, которые посещают данный торговый центр



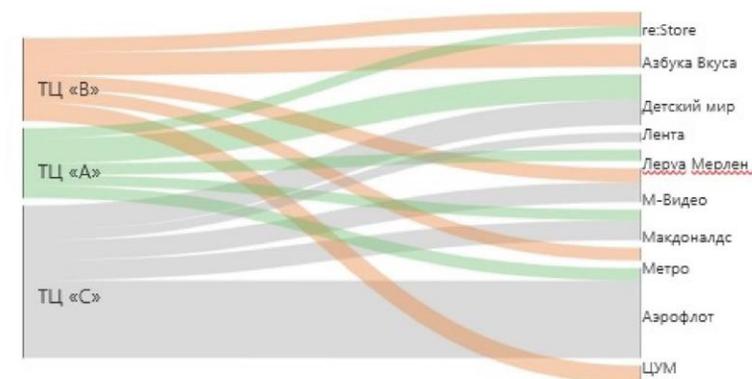
### Клиентская база и пересечение клиентских баз

Число клиентов ТЦ, число уникальных клиентов и клиентов, совершающих покупки в нескольких из исследуемых ТЦ



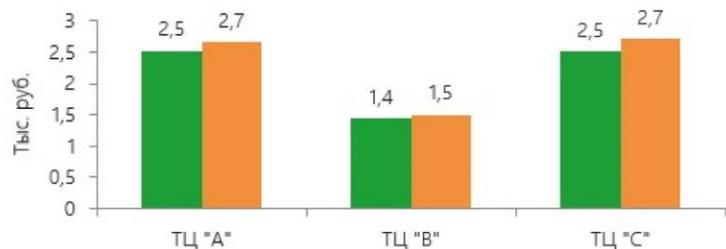
### Любимые бренды

ТОП брендов по суммам трат в ТЦ и вне ТЦ для всей аудитории ТЦ или отдельных категорий



### Средний чек

■ Средний чек ■ Средний чек на визит



### Средние траты клиента ТЦ в год



### Основные категории трат (на примере ТЦ «А»)



# Монетизация данных. Кейсы

## Аналитика города и туристических потоков для развития инфраструктуры туризма



# Монетизация данных. Резюме

- ➔ Монетизация данных дает значимый прирост бизнеса. Например, для финансовых организаций это оценочно составляет до 1,9% дополнительного роста отрасли до 2024 года
- ➔ Монетизация данных может быть внутренней, обеспечивающей повышение эффективности процессов и сервисов, или внешней – направленной на развитие новых направлений бизнеса
- ➔ Существует 4 основных бизнес-модели: подписка, доля прибыли, плата за объем, единовременная оплата
- ➔ Аналитика данных может применяться в самых разных отраслях и сферах деятельности организаций, от строительства и туризма до торговли и добычи полезных ископаемых

# Рынок больших данных

# Развитие больших данных в России способно улучшить качество жизни и принести существенный экономический эффект в 0,3% роста ВВП за 5 лет



## Эффект на существующие отрасли

Увеличение выручки и повышение эффективности таких отраслей, как торговля и добыча полезных ископаемых



## Рост ВВП



Масштабное внедрение больших данных может обеспечить следующий эффект к 2024 vs 2019

- › **0,3% роста ВВП до 2024** = ~20% от 1,3% среднегодового исторического роста ВВП<sup>1</sup>
- › **₽1,3 трлн эффекта в 2024** = ~15% выручки от экспорта нефти в 2018



## Развитие новой отрасли

Рост продуктов и сервисов, позволяющих компаниям использовать большие данные для решения своих задач

## Рабочие места



Развитие отрасли больших данных позволит создать **большое количество привлекательных рабочих мест**

## Качество жизни



Внедрение технологий больших данных в такие направления, как **здравоохранение и умные города** позволит улучшить качество жизни населения



**Сбербанк** является соучредителем АБД, имеет представителя в Правлении АБД. Представитель Сбербанка возглавляет стратегический комитет



АССОЦИАЦИЯ  
БОЛЬШИХ ДАННЫХ

**Ассоциация Больших Данных** объединяет экспертизу «Яндекс», VK, «Сбербанк», «Газпромбанк», «ВТБ», «Тинькофф Банк», «МегаФон», «Ростелеком», oneFactor, QIWI, «Билайн», Аналитический центр при Правительстве РФ, Центр стратегических разработок, «Фонд Сколково»

# Глобальные тренды, обуславливающие перспективы роста рынка больших данных, остаются актуальными и для РФ

## Объем глобального рынка больших данных и бизнес-аналитики



- По различным оценкам объем рынка БД может достигнуть от 163 до 512 млрд долл. США в 2026 г., при этом среднегодовой темп роста может достигать от 11 до 14,8%

## Глобальный рынок больших данных

### Драйверы

- Рост внедрения программного обеспечения для анализа больших данных компаниями, популяризация data-driven подхода в управлении бизнесом
- Рост спроса на облачное программное обеспечение для анализа больших данных среди малого и среднего бизнеса
- Возросшая доступность центров обработки данных, предоставляемых поставщиками облачных технологий

### Барьеры

- Высокая стоимость внедрения БД решений за счет затрат как на технологию, так и на смежные услуги (консалтинг, техобслуживание, установка, интеграция, модернизация и другие)
- Нехватка квалифицированных рабочих кадров

# Эффект от внедрения технологий больших данных по отраслям экономики

## Базовый сценарий

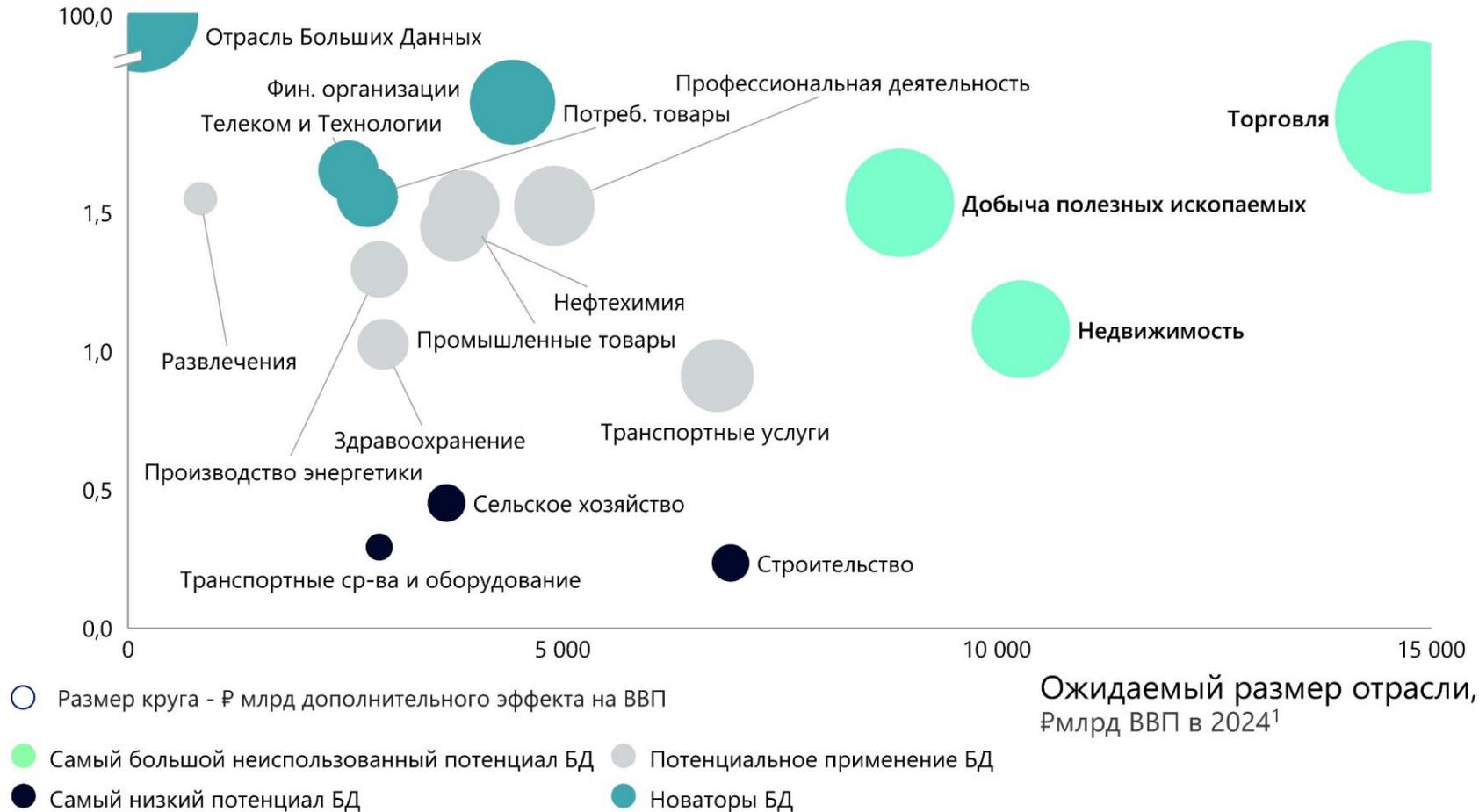
	Прогноз ВВП на 2024 г., млрд руб. <sup>1 23</sup>	Доля в 2024 г. по сравнению с 2019 г., % от ВВП <sup>2</sup>
Торговля	14 778	1,8
Недвижимость	10 277	1,1
Горнодобывающая отрасль	8 883	1,5
Строительство	6 936	0,2
Транспорт	6 782	0,9
Прочие	6 513	1,0
Профессиональные услуги	4 906	1,6
<b>Финансовые учреждения</b>	<b>4 425</b>	<b>1,9</b>
Промышленные товары	3 865	1,5
Нефтехимическое производство	3 756	1,5
Сельское хозяйство	3 665	0,5
Здравоохранение	2 935	1,0
Транспортные средства и оборудование	2 891	0,3
Энергетика	2 889	1,3
Потребительские товары	2 754	1,6
<b>Телекоммуникации и технологии</b>	<b>2 536</b>	<b>1,7</b>
Развлечения	833	1,6
Государственный сектор	7 770	н/д
<b>Всего</b>	<b>97 393</b>	<b>1,2</b>

В 2024 году эффект от внедрения продуктов и технологий больших данных увеличится на 1,2% как доля от ВВП

1. Базовый сценарий Минэкономразвития в постоянных ценах; 2. анализ BCG на основе практического опыта, интервью с экспертами, открытых источников; 3. цветовая кодировка на основе текущего состояния отрасли на 2019 г.;

# Большие данные в торговле, добыче полезных ископаемых и недвижимости обладают наибольшим потенциальным эффектом на ВВП

Эффект от Больших данных  
% от отраслевого ВВП в 2024



## ПРИМЕНЕНИЯ БД

### Торговля

- Оптимизация ассортимента
- Оптимизация цепочки поставок
- Прогнозирование спроса
- Гипер-персонализированное предложение
- Продвинутое таргетирование трафика

### Добыча полезных ископаемых

- Оптимизация продуктового ассортимента
- Оптимизация производственной линии с помощью 'цифровых двойников'
- Управление снабжением в реальном времени

### Недвижимость

- Продвинутое таргетирование клиентов
- Умные системы домоуправления
- Динамическое ценообразование

# Чтобы раскрыть потенциал больших данных России необходимо преодолеть пять групп барьеров

Барьер	Описание	Иллюстративные примеры	Роль профессионального сообщества
 <b>Таланты</b>	Нехватка специалистов, необходимых для массового внедрения Больших данных	Более 6 тыс. открытых вакансий в России на текущий момент <sup>1</sup>	Подтвердить важность этих факторов в рамках программы “Цифровая Экономика” и осуществлять мониторинг выполнения ключевых задач
 <b>Инфраструктура</b>	Ограниченное предложение недорогих масштабируемых ресурсов хранения и обработки	Мощность установленных ЦОДов в России ~85 тыс. кв. м. против ~65 млн кв. м. в США <sup>2</sup>	
 <b>Доступность данных</b>	Ограниченная способность получать доступ, обрабатывать и обмениваться различными типами данных	Барьеры для обработки и распространения личных данных Отсутствие стандартизации и регулярного обмена отраслевых и государственных данных	Обосновать необходимость изменения текущей ситуации + Вести диалог с ключевыми ЛПР + Продвигать устранение барьеров
 <b>Исследование и Идеи</b>	Сложности в исследовании и инновациях в области Больших данных	Отсутствие необходимой среды и ресурсов, необходимых для исследования применения существующих технологий и разработки новых	
 <b>Масштабирование</b>	Невозможность массово имплементировать Большие данные в масштабах всей экономики	Отсутствие широкомасштабного внедрения Больших данных в традиционных отраслях и целевой поддержки продуктов на основе Больших данных в инновационных отраслях	

Барьеры определены/детализированы с использованием 3 линз:

- 1 государственные стратегии Больших данных,
- 2 мнения зарубежных экспертов по Большим данным в СНГ,
- 3 обзор связанных тем в соответствующих дорожных картах программы “Цифровая Экономика”

1. HeadHunter; Большие Данные, искусственный интеллект и машинное обучение; по состоянию на 11/07/2019;  
2. IDC, iKS-Consulting

# Профессиональное сообщество продвигает 6 больших инициатив в целях развития рынка больших данных

## Умеренный эффект

## Существенный эффект



### Доступность данных

#### 1 Упрощенный доступ и обработка данных

Позволить пользователям одновременно и дистанционно давать согласие на несколько целей использования их данных

Например, согласие может быть дано онлайн

Позволить компаниям обрабатывать персональные данные для широкого круга целей при соблюдении определенных требований

Например, с использованием исключительно машинной обработки

Запустить массовую государственную цифровизацию в областях, релевантных для Больших данных, с фокусом на стандартизацию данных

Например, принадлежащие государству данные о здоровье людей



#### 2 Обеспечение возможности обмена/обогащения данных

Позволить игрокам делиться анонимными персональными данными на коммерческой основе

Например, передача анонимных телеком данных ритейлерам

Поощрять обмен отраслевыми данными внутри и между отраслями через саморегулируемые стандарты

Например, обмен эксплуатационных данных доменных печей

Позволить государству делиться определенными типами релевантных данных с частным сектором

Например, аудио или видео с камер наблюдения



### Исследование и Идеи

#### 3 R&D песочницы для исследования Больших данных

Определить законом контролируруемую среду экспериментирования с ослабленным регулированием

Например, облегченные ограничения на использование персональных данных

Обеспечить вовлечение регуляторов для оптимизации одобрений при последующем крупномасштабном развертывании

Например, вовлечение Министерства Цифрового Развития

Обеспечить "озера данных" со стандартизированными данными и технологические библиотеки

Например, заранее обученные алгоритмы машинного обучения



#### 4 Финансирование инноваций и ресурсная экосистема

Обеспечить инновационные команды выделенным доступом к "озерам данных" с труднодоступной отраслевой информацией

Например, эксплуатационные данные сетей электроснабжения

Оптимизировать процессы бэк-офисного типа путем предоставления доступа к юристам, бухгалтерам и специалистам по патентам

Например, предоставление компаниями части своих ресурсов

Внедрить инвестиционную платформу, соединяющую квалифицированных инвесторов с отобранными инициативами

Например, внедрение механизмов долевого владения



### Масштабирование

#### 5 Стратегии Больших данных традиционных индустрий

Создать стандарт для внедрения Больших данных в компании с государственным участием

Например, в таких отраслях как добыча полезных ископаемых, промышленные товары

Ввести ориентированные на результат стимулы для компаний частного сектора стимулы

Например, для достижения целей по росту выручки

Создать проектный и технический кадровый резерв, чтобы помочь компаниям внедрять Большие данные и обучать их команды

Например, выделение времени существующих команд



#### 6 Внутренние стимулы для инновационных отраслей

Внедрить упрощенный процесс получения необходимых сертификатов и патентов для продуктов и услуг на основе Больших данных

Например, более простой процесс сертификации

Устранить выборочные барьеры для экспорта продуктов и сервисов, построенных на технологиях Больших данных

Например, освободить компании от пошлин

Провести кампании по повышению осведомленности об экспортном потенциале продуктов на технологиях Больших данных

Например, со стартапами вне Москвы

# Вклад Больших данных в ВВП может вырасти на 1,2%-1,8% к 2024 против 2019 при реализации 6 больших инициатив

Обязательное условие для преодоления "Бездействия": комплексные меры в части инфраструктуры и талантов



# Технологические ограничения, на фоне ухода из России крупных международных игроков, сокращают предложение на рынке больших данных

## Факторы, влияющие на рынок БД

## Примеры ограничений (санкций)<sup>1</sup>

## Текущее положение



### Таланты

- Отказ от работы на российском рынке крупных международных вендоров
- Сокращение и релокация российских офисов и центров разработки международных ИТ-компаний

- Увеличение разрыва между потребностями рынка и реальным объемом квалифицированных кадров
- Рост количества вакансий в ИТ-секторе
- По оценкам РАЭК, в апреле 2022 года из России уехали 70-100 тыс. ИТ-кадров



### Инфраструктура

- Запрет на экспорт в Россию высокотехнологичного оборудования и микроэлектроники, ПО и решений на базе ИИ
- Запрет на экспорт в Россию микросхем, полупроводников
- Отказ от работы на российском рынке крупных международных облачных сервисов
- Запрет оказание технологических услуг локальным вендорам, включая участие в разработках и использование архитектуры

- Кратный рост стоимости оборудования и услуг поддержки вендоров
- Снижение инвестиций в основной капитал, по оценке экспертов ЦМАКП<sup>1</sup>, в 2022 г. может достичь 25-30%
- Ограничен доступ к инструментам работы с БД, включая доступ к специализированному ПО
- Ограничение поддержки, отзыв лицензий и сертификатов международных вендоров
- Рост дефицита мощности ЦОД в России



### Исследования и идеи

- Приостановка совместных научных программ с исследовательскими организациями
- Ограничение доступа к результатам международных научных исследований и информационным ресурсам

- Снижение предложения и ухудшение качества прогнозирования в связи с ухудшением доступности данных
- Сокращение спроса компаний на услуги, связанные с обработкой данных в рамках НИОКР

<sup>1</sup> – НК Центр макроэкономического анализа и краткосрочного прогнозирования.

# Отраслевые санкции дополнительно ограничивают масштабирование решений на основе больших данных

на примере финансового сектора

Примеры ограничений (санкций)<sup>1</sup>



**Санкции против кредитных организаций (банков), страховых и лизинговых компаний**

Влияние на применение технологий больших данных организациями финансового сектора

- Сокращение спроса на решения в сфере Больших Данных (уровни сбора, обработки и хранения данных) из-за снижения физической доступности оборудования и ПО, высокой стоимости параллельного импорта (50-200 раз)
- Усложнение масштабирования решений на основе больших данных в рамках автоматизации процессов (риск-менеджмент, прогностическое моделирование, скоринг, управление киберриском, углубленная сегментация клиентов, управление активами и пр.)



**Приостановка вендорской поддержки и ограничение доступа к аналитическим финансовым продуктам**

- Снижение запроса финансовых организаций на применение вертикальных решений и услуг, включая использование облачных решений в рамках основных операционных процессов
- Сокращение объемов и качества обработки и хранения данных в связи с повышением рисков общей работоспособности и безопасности основных процессов
- Ограничение доступа к внешним данным, связанное с отключением терминалов, закрытия доступа к ленте данных и на электронные торговые платформы ведущих информационных агрегаторов (например, Bloomberg, Reuters)



**Приостановка работы глобальных платежных систем Visa, MasterCard, American Express, JCB**

- Сокращение транзакционного бизнеса и отказ от инвестиций во внедрение и масштабирование решений для автоматизации аналитики и проведения платежей с использованием больших данных

1 – неисчерпывающий список. 2 – экспертная оценка

# Актуальные вызовы рынка больших данных

## ТЕНДЕНЦИИ

- 1 Ожидаемый значительный спад в экономике и в ключевых отраслях в 2022-2024 гг.
- 2 Введение санкций на импорт и использование технологий и услуг и на международное сотрудничество, затрагивающих рынок больших данных и создающих риски для национальной технологической инфраструктуры
- 3 Технологические и операционные ограничения, уход из России крупнейших мировых игроков на многих рынках, включая рынок больших данных
- 4 Значительный отток из российской экономики специалистов в сфере работы с данными

## ПОСЛЕДСТВИЯ



Сжатие спроса российской экономики на вертикальные решения и услуги в сфере больших данных



Дополнительное ограничение спроса на использование технологий и услуг в связи с геополитической обстановкой, включая санкции



Значительное сокращение предложения вертикальных решений, услуг и технологических инструментов на российском рынке



Необходимость выстраивания новой системы взаимодействия участников российского рынка больших данных, в т.ч. с зарубежными партнёрами

# Рынок больших данных. Резюме

- Развитие больших данных дает существенный вклад в улучшение качества жизни и рост экономики страны (~1,3 трлн рублей в 2024 году). Наибольший потенциал развития - в торговле, добыче полезных ископаемых и недвижимости
- Существует 5 групп барьеров, препятствующих развитию рынка: нехватка специалистов, ограничения инфраструктуры, доступность данных, сложности в развитии инноваций, сложности масштабирования решений
- Для развития рынка больших данных необходимо реализовать 6 больших инициатив:
  - упрощенный доступ и обработка данных
  - обеспечение возможности обмена и обогащения данных
  - R&D песочницы для исследования больших данных
  - финансирование инноваций и ресурсная экосистема
  - стратегии больших данных традиционных индустрий
  - внутренние стимулы для инновационных отраслей
- Актуальность развития рынка больших данных и снятия барьеров усиливается на фоне международных ограничений. Необходимо выстраивать новую систему взаимодействия участников рынка и развивать собственных технологические решения

# Благодарю за внимание!

