

# **Инвариантная самостоятельная работа 4**

## **Базы данных мира**

### **Оглавление**

Оглавление.....	1
Что такое база данных?.....	2
Что такое банк данных?.....	3
В чем отличие базы данных от банка данных?.....	4
Что такое Big Data?.....	5
Что такое СУБД?.....	6
Что такое аналитика больших данных?.....	7
Поставщики Big Data в России.....	8
Поставщики Big Data в мире.....	9
Amazon Web Services.....	9
Google Cloud.....	9
Microsoft Azure.....	9
IBM.....	9
Oracle.....	9
Alibaba Cloud.....	10
Технические характеристики баз данных.....	11
Использованные ресурсы.....	12

## **Что такое база данных?**

База данных — это организованная совокупность данных, предназначенная для хранения, поиска и обработки с использованием специального программного обеспечения. Данные структурированы по определённым правилам и связаны между собой.

## **Что такое банк данных?**

Банк данных — это более широкое понятие, включающее в себя базу данных, а также технические, программные и организационные средства для её ведения, хранения, поиска и предоставления информации пользователям.

## **В чем отличие базы данных от банка данных?**

База данных — это часть банка данных, а банк данных — это полноценная система, объединяющая информацию и средства для её использования.

## **Что такое Big Data?**

Big Data (большие данные) — это массивы данных большого объёма и разнообразия, поступающие с высокой скоростью и требующие особых технологий обработки, хранения и анализа. Обычно описываются по принципу "5V": Volume, Variety, Velocity, Veracity, Value.

## **Что такое СУБД?**

СУБД (система управления базами данных) — это программное обеспечение, предназначенное для создания, управления, модификации и обеспечения доступа к базам данных. Примеры: MySQL, PostgreSQL, Oracle Database, SQLite.

## **Что такое аналитика больших данных?**

Аналитика больших данных — это процесс обработки и анализа больших массивов информации для выявления закономерностей, трендов, прогнозов и поддержки принятия решений. Используются методы машинного обучения, статистики, визуализации и др.

## Поставщики Big Data в России

№	Компания	Примеры решений	Режим доступа
1	Яндекс	Yandex DataLens, Метрика	<a href="https://datalens.tech/">https://datalens.tech/</a> <a href="https://metrika.yandex.ru">https://metrika.yandex.ru</a>
2	Сбер	SberAnalytics	<a href="https://sberanalytics.ru/">https://sberanalytics.ru/</a>
3	VK	VK Cloud	<a href="https://cloud.vk.com/">https://cloud.vk.com/</a>
4	МТС	МТС AI	<a href="https://mts.ai/solutions/">https://mts.ai/solutions/</a>
5	Тинькофф	Собственная инфраструктура	<a href="https://tinkoff-group.com/">https://tinkoff-group.com/</a>



## **Поставщики Big Data в мире**

### **Amazon Web Services**

Amazon Web Services (AWS) — американская компания, предоставляющая широкий спектр облачных сервисов. В области Big Data популярны такие решения, как Amazon Redshift (аналитическая СУБД) и AWS Glue (инструмент для подготовки и трансформации данных).

### **Google Cloud**

Google Cloud — ещё один крупный игрок из США, предлагающий мощные инструменты для хранения и анализа данных. Основные решения — BigQuery (облачная аналитическая платформа) и Dataflow (сервис потоковой обработки данных).

### **Microsoft Azure**

Microsoft Azure - облачная платформа от Microsoft. Среди её решений в сфере Big Data - Azure Synapse Analytics (интеграция хранилища и аналитики) и Azure Data Lake (масштабируемое хранилище для больших данных).

### **IBM**

IBM — американская компания, работающая в сфере ИТ и аналитики. Её решения включают IBM Watson (платформа ИИ) и IBM Cloud Pak for Data (интегрированная платформа для анализа больших данных).

### **Oracle**

Oracle — известный поставщик СУБД, также активно развивает Big Data-направление. Основной продукт — Oracle Big Data Service,

предлагающий инструменты для хранения, обработки и анализа информации.

### **Alibaba Cloud**

Alibaba Cloud — крупнейший поставщик облачных решений в Китае. В его портфеле по Big Data выделяются MaxCompute (сервис обработки данных) и DataWorks (платформа управления данными).

## **Технические характеристики баз данных**

Базы данных имеют следующие основные технические характеристики:

- Объём хранения может варьироваться от нескольких гигабайт до нескольких петабайт, особенно в случае систем, работающих с Big Data.
- Модель данных определяет способ структурирования информации. Наиболее распространённые модели: реляционная (таблицы), документная (например, JSON), графовая (узлы и связи) и колонковая (оптимизирована под аналитику).
- Производительность зависит от архитектуры системы, наличия индексов, кэширования и оптимизации запросов. Она критически важна для обработки больших объёмов данных.
- Масштабируемость может быть вертикальной (увеличение мощности одного сервера) или горизонтальной (добавление новых узлов в кластер).
- Доступность характеризует устойчивость к сбоям и отказам. Используются технологии резервного копирования, репликации и кластеризации для обеспечения непрерывной работы.
- Безопасность включает механизмы аутентификации, авторизации, шифрования данных и ведения журналов действий пользователей.
- Поддержка запросов реализуется через языки и интерфейсы доступа, такие как SQL (для реляционных БД), NoSQL API (для нереляционных), а также REST и другие веб-интерфейсы.

## **Использованные ресурсы**

1. Кузнецов, С. М. Базы данных: учебное пособие. — М.: Академия, 2020. — 224 с.
2. Тураев, В. А. Системы управления базами данных. — СПб.: Питер, 2019. — 320 с.
3. Oracle. Oracle Big Data Solutions [Электронный ресурс]. — Режим доступа: <https://www.oracle.com/big-data/>
4. Amazon Web Services. Big Data on AWS [Электронный ресурс]. — Режим доступа: <https://aws.amazon.com/big-data/>