**Задание 7. Крупнейшие базы данных мира.**

Задание.

Ответьте на вопросы.

Что такое база данных?

совокупность данных, хранимых в соответствии со схемой данных, манипулирование которыми выполняют в соответствии с правилами средств моделирования данных (ГОСТ Р ИСО МЭК ТО 10032-2007)

Что такое банк данных?

Банк данных – это специальным образом организованная система данных (баз данных), технических, языковых, программных и организационно-методических средств, которые предназначены обеспечить централизованное накопление и коллективное многоцелевое использование данных.

Что такое Big Data?

Это обозначение структурированных и неструктурированных данных огромных объёмов и значительного многообразия, эффективно обрабатываемых горизонтально масштабируемыми программными инструментами, появившимися в конце 2000-х годов и альтернативных традиционным системам управления базами данных и решениям класса Business Intelligence

Что такое СУБД?

Система управления базами данных, сокр. СУБД (англ. Database Management System, сокр. DBMS) — совокупность программных и лингвистических средств общего или специального назначения, обеспечивающих управление созданием и использованием баз данных. Комплекс программ, позволяющих создать базу данных (БД) и манипулировать данными (вставлять, обновлять, удалять и выбирать). Система обеспечивает безопасность, надёжность хранения и целостность данных, а также предоставляет средства для администрирования БД.

Что такое аналитика больших данных?

Это применение передовых методов анализа к очень большим и разнообразным наборам данных, включая структурированные, частично структурированные и неструктурированные данные из различных источников в масштабах от терабайт до зеттабайт.

Какие компании являются поставщиками Big Data в России?

Основные поставщики больших данных в России — поисковые системы. Они имеют доступ к массивам данных, а кроме того, обладают достаточной технологической базой для создания новых сервисов. Например: Яндекс, Mail.ru group, Rambler.

Какие компании являются поставщиками Big Data в мире?

IBM, Microsoft, Oracle, Sap и другие.

Технические характеристики баз данных.

Максимальный размер таблиц в MySQL 3.22 до 4 гигабайт, в последующих версиях ограничений нет.

Размер таблицы ограничен её типом. В общем случае тип MyISAM ограничен предельным размером файла в файловой системе операционной системы. Например, в NTFS этот размер теоретически может быть до 32 эксабайт. В случае InnoDB одна таблица может храниться в нескольких файлах, представляющих единое табличное пространство. Размер последнего может достигать 64 терабайт.

В отличие от MyISAM, в InnoDB имеется значительное ограничение на количество столбцов, которое можно добавить в одну таблицу. Размер страницы памяти по умолчанию составляет 16 килобайт, из которых под данные отведено 8123 байта. Размер указателя на динамические поля составляет 20 байт. Таким образом, в случае использования динамического формата строки (ROW\_FORMAT=DYNAMIC), одна таблица может вместить максимум 409 столбцов типа blob или text.

Какая из баз данных наиболее удобна в использовании и содержит данные по профилю вашей подготовки?

Amazon Aurora ([ссылка на сайт](https://aws.amazon.com/ru/rds/aurora/?hp=tile&so-exp=below&aurora-whats-new.sort-by=item.additionalFields.postDateTime&aurora-whats-new.sort-order=desc))

Кто ее разработчике и/или поставщик Big Data.

BryteFlow Blend, Fivetran, Information Builders, Amazon Web Services и др.

Технические характеристики базы данных.

* Высокая производительность и масштабируемость (для масштабирования ресурсов и производительности чтения можно добавлять до 15 реплик чтения с низкой задержкой в трех зонах доступности. При необходимости Amazon Aurora автоматически увеличивает объем хранилища вплоть до 128 ТБ)
* Высокая доступность и надежность (Amazon Aurora обеспечивает доступность на уровне более чем 99,99 % с поддержкой шести копий данных в трех зонах доступности и непрерывно выполняемым резервным копированием ваших данных в хранилище Amazon S3)
* Высокий уровень безопасности (Amazon Aurora обеспечивает несколько уровней защиты базы данных. К ним можно отнести сетевую изоляцию с помощью Amazon VPC, шифрование данных в состоянии покоя с использованием ключей, которые вы создаете и контролируете в сервисе AWS Key Management Service (KMS), а также шифрование передаваемых данных с использованием SSL)
* Совместимость с MySQL и PostgreSQL (ядро базы данных Amazon Aurora полностью совместимо с существующими базами данных с открытым исходным кодом MySQL и PostgreSQL; совместимость с новыми выпусками регулярно добавляется)
* Полностью управляемая БД (Amazon Aurora работает под полным управлением сервиса Amazon Relational Database Service (RDS). При работе с базами данных больше не требуется заниматься такими административными задачами, как выделение необходимого оборудования, установка исправлений ПО, установка и настройка самой базы данных, а также ее резервное копирование. Aurora автоматически и непрерывно выполняет мониторинг базы данных)
* Поддержка миграции (совместимость с MySQL и PostgreSQL обеспечивает привлекательность Amazon Aurora в качестве целевой БД при миграции баз данных в облако. Если требуется выполнить миграцию баз данных MySQL или PostgreSQL, в документации по миграции приводится список инструментов и вариантов миграции).