

Темы для ИДЗ

Темы размечены по сложности (простые и сложные). Для простых тем дедлайны раньше, но задания проще. Для сложных тем задания сложнее, но и дедлайны позднее.

Помимо задания, для каждой ИДЗ требуется реализовать **массовый импорт-экспорт данных**.

Темы 2026

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий	
1	ИС для театральных декораций	Сложная	MongoDB	Синтетические данные	Задача - сделать веб-приложение для хранения, поиска и редактирования информации о театральных декорациях с учетом их специфики.	
2	БД актеров	Простая	MongoDB	Синтетические данные. Для понимания сути проблемы - посмотрите в открытом доступе видео визитки актеров и их страницы на сайте театров.	Задача - создать каталог профилей актеров со всевозможной дополнительной информацией + поиск + редактирование.	
3	ИС для ателье	Сложная	MongoDB	Синтетические данные	Задача - создать веб-приложение для ателье, где будут роли заказчиков, работников, руководства. В приложение должна быть возможность создать заказ, трекать его выполнение, хранить всю промежуточную информацию, подводить статистику, считать финансы.	
4	Хранилище изображений с CRUD	Простая	MongoDB	Синтетические данные	Задача - реализуйте веб-приложение с REST интерфейсом для CRUD изображений. В веб-интерфейсе реализуйте возможность поиска изображений по метаданным, датам, данным самого изображения. Храните все логи.	
5	БД биометрии СКУД для университета (голос, лица, логи, доступы)	Сложная	MongoDB	Синтетические данные	Задача - реализовать веб-приложение для хранения биометрии, для контроля доступа и логирования всех связанных событий.	
6	Система записи на самостоятельную работу в комп классе	Простая	ArangoDB	Синтетические данные	Реализуйте приложение для записи на самостоятельную работу в один из Н компьютерных классов.	
7	БД для наблюдений за животными	Простая	Neo4j	Синтетические данные	Задача - создать веб-приложение, в котором можно делиться своими наблюдениями за животными (фото, видео, отметки на карте, лайки, комментарии).	
8	ИС анализа диссертаций в открытом доступе	Сложная	ArangoDB	https://vak.gisnauka.ru/adverts-list/advert	Задача - реализовать приложение, которое принимает на вход тексты диссертаций и авторефератов, а также метаданные о защите / авторе / специальности и тд. Приложение должно парсить тексты, разделяя их на характерные компоненты (элементы введения, список литературы, размеры глав и тд). Предоставлять поиск и стат анализ по данным	
9	Система для распределения студентов по практикам	Простая	Neo4j	ЛК партнера	Задача - давайте сделаем ЛК-партнера 2.0 с очень умными рекомендациями, навыками (и их подтверждением) студентов, структурированными вакансиями (практиками), аналитикой и поиском.	
10	Бэкенд для gprs mtoorg	Сложная	Neo4j	МН	Давайте сделаем наш ответ Mystery Hike (но с упором на бэкенд и БД).	
11	ИС для бесконечных крестиков-ноликов.	Сложная	MongoDB	Синтетические данные	Сделать платформу для запуска пользовательских программ-ботов, сохранять сами решения, логи запусков, состояние поля, подводить статистику.	
12	Веб-приложение для дыхательных упражнений.	Простая	ArangoDb		Хранить каталог упражнений, отдельные карточки и визуализацию каждого. Режим выполнения упражнения - визуализировать все фазы. Пользователи, комментарии, статистика, отзывы, профили.	
13	Каталог персонажей вселенной Властелина колец	Простая	Neo4j		Хранить в виде графа данные по всем персонажам, предметам, временам, локациям и тд из вселенной ВК. Предоставлять поиск, аналитику, агрегацию данных	
14	Построитель маршрутов для раскраски карты города.	Сложная	Neo4j	МН	В некоторых GPS MMORPG вы закрашиваете карту, просто перемещаясь по городу (вы на карте представляете круг фиксированного диаметра). Задача - сделать веб-приложение, которое принимает на вход открытые данные OSM и растровую карту текущей закрашки. На выходе необходимо построить несколько адекватных маршрутов, чтобы минимумом усилий закрасить область.	
15	ИС для симуляции уборки снега.	Сложная	Neo4j	OSM	В интерфейсе приложения задаются координаты стоянок уборочной техники, снегоплавильных полигонов, станций тех обслуживания. В качестве карты выступают реальные OSM данные Санкт-Петербурга. Приложение проводит расчет лучших маршрутов для техники с учетом различных ограничений (пробки и доступность улиц, ресурс и вероятность поломки техники), считает статистику. Приложение хранит результаты моделирования, расстановку объектов на картах и другую информацию.	
16	Анализатор графов вызова для программ на С.	Сложная	Neo4j	Синтетические данные	Необходимо составить теоретический граф (с точки зрения исходника) - какие функции будут и как вызваны, сопоставить его с фактическим (запустив программу в callgrind). Провести анализ получившегося совпадения.	

17	Анализатор отчетов в досх.	Сложная	Neo4j	Синтетические данные	Необходимо сделать веб приложение, которое анализирует учебные отчеты с точки зрения структуры (превращает их в деревья предложений, слов), а также вычисляет различные статистики, находит между разными работами пересечения.	
18	Инструмент для кластеризации данных протектинга.	Сложная	MongoDB	Синтетические данные (надо записать самому)	На входе у нас логи moodle, списки действий пользователей на веб странице, статистика распознавания лиц / OCR на скринкастах Сервис должен уметь хранить эти разнородные наборы, объединять их по признакам сессий и принадлежности к студенту, а также фильтровать различными способами, агрегировать. Самое главное - должны быть реализованы алгоритмы кластеризации, чтобы строить и визуализировать кластеры среди сессий, находить аномальные.	
19	Инструмент для кластеризации логов apache	Простая	MongoDB	Синтетические данные	Задача - создать систему для хранения логов Apache (всех видов логов), поиска, группировки и обработки + применение различных методов кластеризации (с сохранением в БД данных каждой попытки кластеризации).	
20	Веб приложение для управления группой наземных роботов.	Сложная	MongoDB	Синтетические данные	Нужно сделать приложение, решающее задачу хранения информации о N группах роботов. Храним как метаданные самой группы, так и метаданные конкретных роботов, их координаты, траектории, бинарные данные (кадры с камеры), храним текущие задания для группы, логи. Визуализируем положение на карте.	
21	Каталог ВУЗов для поступающих - специальности, факультеты, условия, баллы, калькуляторы поступления и аналитика	Простая	MongoDB	etu.ru	Готовим наш ответ вузопедии.	
22	Родословное древо дворянских фамилий / царей / королей	Простая	Neo4j	https://imperialhouse.ru/rus/imperialhouse/succession/tree.html	Делаем визуализатор и анализатор родословных деревьев для дворян и правителей.	
23	Сервис для электронных таблиц и форм	Сложная	ArangoDB	Синтетические данные	Давайте задумаемся на создание нашего ответа GoogleSheets	Forms. Реализуем в простом виде создание / представление / редактирование / анализ таблиц и форм.
24	Приложение-стенд для производительных (и продвинутых) 3d-индексов - хранение и обработка облаков точек	Сложная	MongoDB	Синтетические данные	Задача - сделать веб приложение, которое хранит данные о тестовых облаках точек, разных алгоритмах построения индексов, запусках экспериментов, а также позволяет все это анализировать, искать, агрегировать.	
25	CRM для кадровиков - вакансии, отклики, кандидаты, задания, собеседования	Простая	Neo4j	Синтетические данные	Задача - сделать веб-приложение для кадровиков, которое поможет в найме сотрудников (обработка откликов, соискатели, собеседования и их статусы, кто когда и что спросил, личные данные ...).	
26	Простые отметки на карте - грибные места, координаты участков, отметки на карте, маршруты	Простая	MongoDB	OSM	Задача - реализовать веб-приложение, в котором можно делать простые заметки на карте. Отмечать интересные места, зоны, маршруты и тд. Искать и анализировать их, соотносить с объектами на карте.	
27	Сервис по вывозу мусора (Приложение одновременно для курьеров и клиентов)	Простая	ArangoDB	OSM	Задача - реализовать веб-приложение с ролями курьер, клиент, администратор, где был бы реализован сервис по вывозу бытового мусора. Трекинг событий, финансов, действий, аналитика.	
28	Тренажер для планирования задач в разработке ПО	Сложная	Neo4j	Синтетические данные	на входе задачи в беспорядке с оценками времени и примерными отношениями порядка, на входе надо составить граф выполнения, прикинуть оценки времени по срокам + учесть табель календарь	
29	Веб-приложение для игры hotseat, хранения партий и тестирования ботов для круговых шахмат	Простая	MongoDB	Синтетические данные	Задача - веб-приложение для игры в круговые шахматы. Хранение партий, статистики, аккаунты, подключение ботов и их тестирование.	
30	Каталог народных примет и их накладка на архив наблюдений погоды	Сложная	ArangoDB	Новости Gismeteo?	На сайте Gismeteo каждый день публикуют новость про дату в народном календаре и там каждый раз есть различные приметы, некоторые из которых можно верифицировать. Задача - сделать веб приложение, которое собирает подобные приметы (и структурирует их), собирает погодные наблюдения, составляет отчеты по проверке и адекватности той или другой приметы.	

31	Анализ коммитов репозитория пакетов Linux.	Сложная	MongoDB	Repology и данные git	В repology есть данные, но они не очень хорошие. У некоторых дистрибутивов все пакеты лежат в одном репозитории - Homebrew (core, cask), nixpkgs, Guix, Gentoo, Void, FreeBSD, AUR (https://github.com/archlinux/aur). У некоторых - по одному git-репозиторию на пакет, давайте проанализируем Arch (https://gitlab.archlinux.org/archlinux/packages), Debian (https://salsa.debian.org/debian/), Alt Linux (https://git.altlinux.org/gears/). Там есть защита от парсинга, я думаю, надо попробовать взять список проектов из дампа repology (https://dumps.repology.org/) и клонировать все репозитории по одному. Учтите, что репозитория много 😊 Нужно сделать общую базу всех коммитов и статистику - кол-во уникальных разработчиков в месяц, динамика общего кол-во пакетов, пересечение между разработчиками.	
32	Информационная система центра полиграфических услуг	Простой	Neo4j	Синтетические данные	Задача - сделать сервис для обеспечения работы центра полиграфии (печать, сканирование, ризография). Роли: клиенты, сотрудники, администраторы.	
33	Агрегатор информации о микроклимате в большом помещении (см пример с занятия по Influx - мониторинг микроклимата в ВУЗе)	Сложный	InfluxDb	etu.ru	Задача - сделать сервис, где будут собираться и анализироваться данные о микроклимате для большой группы строений со сложной структурой (на примере ЛЭТИ). Нужны визуализации как по отдельным помещениям, так и по этажам, корпусам, а также страницы самих отдельных единиц, страницы статистики.	
34	Информационная система компании по производству и монтажу кухонных гарнитуров	Сложный	MongoDB	Синтетические данные	Задача - сделать сервис для бизнеса, который занимается уборкой. Основной акцент - прототип системы управления производством (заказы, раскрой, рабочие часы, логистика).	
35	Сервис выбора земельных участков для купли-продажи	Сложный	MongoDB	https://egrp365.org/map/?id=g3c6CO , avito.ru	Задача - создать сервис, в котором можно просматривать и визуализировать объявления о купле-продаже участков. Для участков необходимо составлять аналитику с точки зрения их привлекательности или наоборот - негативных факторов для покупки.	

Своя тема

1. **Интересный, содержательный и нужный** пользователям сценарий использования
2. Должен быть понятный, достаточный и доступный **источник данных** - публичный датасет, публичное API доступа к хранилищу данных
3. Должно присутствовать решение задач **хранения/анализа/представления/агрегации/импорта и экспорта** данных
4. Дополнительный плюс, если есть **связь с существующим проектом** и/или бакалаврской работой
5. Информация об **аналогах**

Источники данных

Помимо ранее упомянутых датасетов, можно также использовать для своей темы:

1. [Список SPARQL-endpoint](#)
2. <https://github.com/blog/2298-github-data-ready-for-you-to-explore>
3. <http://www.foodpedia.tk/about?lang=ru>
4. <http://www.imdb.com/interfaces>
5. https://www.wikidata.org/wiki/Wikidata:Main_Page
6. Выгрузка данных OSM
 1. <https://www.openstreetmap.org/search?query=saint-petersburg#map=10/59.9471/30.2495>
7. [Kaggle](#) - каталог наборов данных
8. [Datahub](#) - каталог наборов данных

From:
<https://se.moevm.pro/> - **МОЭВМ Вики** [se.moevm.pro]

Permanent link:
https://se.moevm.pro/doku.php/staff:courses:no_sql_introduction:course_work:topics

Last update:

