

Темы для ИДЗ

Темы (с 2025 года) 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 2024

№	Название	Предлагаемая СУБД	Набор данных/API	Комментарий
01	Каталог упражнений йоги	MongoDB	Синтетические данные	Задача - сделать каталог упражнений йоги, с подробным описанием, картинками, классификацией, показаниями и противопоказаниями, отзывами и комментариями. Кроме самого каталога, нужен также многокритериальный поиск по упражнениям и комментариям.
02	Сервис поиска психологов	MongoDB	Синтетические данные	Задача - делаем свой аналог b17. Карточки психологов, модераторов и клиентов. Поиск пользователей, отзывы, интерфейсы модерирования, комментарии.
03	Сервис истории домов СПб	ArangoDB	Citywalls	Задача - готовим наш вариант Citywalls с большим упором на классификацию, выделение кластеров и визуализацию.
04	Каталог литературных цитат	MongoDB	Классическая литература 19 века :)	Задача - сделать каталог цитат из классики 19 века, для определенности - отечественной. Нужно предоставить карточки цитат, на которых отобразить произведение, автора, персонажа (если это его слова), теги, похожие цитаты. Также нужен поиск по цитатам, авторам, произведениям (многокритериальный).
05	Сервис поиск переименованных в советское время топонимов в СПб	Neo4j	Citywalls	Задача - выделить из открытых данных именно кейсы переименования, дать по ним поиск и возможность визуализации на карте.
06	Каталог и агрегатор отзывов на озера РФ	Neo4j	Открытые данные OpenStreetMap	Задача - буквально сделать каталог озер на территории РФ. Карточки озер, поиск, страницы отзывов, профили пользователей.

№	Название	Предлагаемая СУБД	Набор данных/API	Комментарий
07	Каталог всех дворов-колодцев СПб	Neo4j	Открытые данные OpenStreetMap	Задача - составить каталог всех дворов-колодцев в СПб. Смысл сервиса - дать возможность найти все дворы, а также вести учет персональных достижений (Кто в сколько дворах побывал). Страницы : профиль пользователя, карточка двора, поиск пользователей и дворов, карта.
08	Сервис для организации субботников	ArangoDb	Синтетические данные	Задача - сделать сервис, где можно как самому организовать субботник, так и присоединится к чужому субботнику. В сервисе должны быть помимо стандартных страниц (Профили пользователей, поиск, комментарии, отзывы) также и страницы с персональной и общей статистикой (кто и где участвовал, сколько кв км убрали ...).
09	Приложение для продажи и обмена комнатными растениями	MongoDB	Синтетические данные	Задача - сделать свое avito с базой знаний и фокусом на комнатных растениях.
10	Приложение для ведения домашней библиотеки	MongoDB	Синтетические данные	Представьте, что у вас есть очень много хороших и интересных книг дома, а также еще больше друзей, которые их таскают. Нужно сделать веб сервис, где подобные люди смогут вести свои домашние библиотеки, а также инициировать обмен / как-то фиксировать, кто что и кому отдал.
11	Сервис-сборщик ответов на яндекс-формы и добавление их в таблицы	MongoDB	Синтетические данные	Задача - нужен сервис для выполнения (В большом объеме) утилитарной задачи, а именно, автоматического импорта ответов из яндекс формы в яндекс таблицу (в качестве примера - гугл формы так умеют из коробки). Ваш сервис должен не просто уметь такое делать, но и 1) обеспечивать синхронизацию произвольного количества пар «форма - таблица», 2) вести логи действий (кто и когда какую форму заполнял, как и куда данные пересылались) 3) для пар в системе обеспечивать администрирование.

№	Название	Предлагаемая СУБД	Набор данных/API	Комментарий
12	Сервис для indie групп - поиск замен музыкантов, реп.точки, концерты	ArangoDB	Синтетические данные	Задача - сделать веб-сервис, где indie-группы смогут дружить и сотрудничать. Профили групп, музыкантов, реп.точки, объявления, комментарии, рейтинги, статистика.
13	Сервис анализа данных для выдачи кредитов - Профили клиентов, составление рейтингов, прогнозы	MongoDB	Синтетические данные	Задача - давайте пофантазируем о том, что мы большой банк, который выдает кредиты. Разработайте веб-приложение, в котором будут храниться сведения о клиентах банка, со всеми необходимыми полями для того чтобы принять решение. Реализуйте в простейшей форме алгоритм принятия решений, а также процедуру согласования данных решений. Визуализируйте в сервисе как сотрудников банка, так и клиентов, а также статистику и историю их взаимодействия.
14	Информационная система для пожарных бригад - Вызовы, Бригады, Автомобили, Оборудование, Оценка ущерба, Звонки на горячую линию, Сотрудники (Операторы горячей линии, Пожарные)	Neo4j	Синтетические данные + Открытые данные OpenStreetMap	Задача - создать сервис, где будет храниться вся необходимая информация для работы одного пожарного депо, а также необходимые веб-страницы, чтобы с этими данными работать.
15	Сервис аренды инструментов для ремонта	MongoDB	Синтетические данные	Задача - создать сервис для фирмы, которая сдает в аренду различное строительное оборудование (от молотков до катков и экскаваторов). Нужны интерфейсы как для сотрудников, так и для клиентов.
16	Сервис для Тайного Санты	MongoDB	Синтетические данные	Задача - делаем сервис, где несколько независимых групп пользователей могут играть в тайного Санту.
17	Сервис для организации длинных экскурсионных маршрутов по СПб	Neo4j	Citywalls + Открытые данные OpenStreetMap	Задача - сделать сервис, который сможет составить длинный пешеходный маршрут по замысловатому запросу, а также вести учет пользовательских маршрутов, статистики прохождения этих самых маршрутов. Например, маршрут через все мосты СПб, обход всех площадей центрального района.

№	Название	Предлагаемая СУБД	Набор данных/API	Комментарий
18	Сервис-агрегатор для поиска свободных мест на занятиях йоги	MongoDB	Синтетические данные	Задача - сделать сервис, где клиенты смогут видеть общее расписание нескольких центров йоги, а также записываться в свободные слоты.
19	Сервис управления сетью коммерческих автоматизированных складов	ArangoDb	Синтетические данные	Задача - сделать сервис для бизнеса, который предоставляет услуги аренды ячеек на складах. Нужны интерфейсы для как сотрудников (владелец бизнеса, директор склада, рабочий, системный администратор), так и пользователей, интерфейсы для оплаты на заглушках.
20	Сервис для организации ремонтов в ВУЗе - сроки, материалы, работы, переселение работников в новые кабинеты, оценка рисков	MongoDB	etu.ru	Задача - сделать веб-сервис для планирования хозяйственных работ в большой, протяженной организации, в которой работает много людей. Основной упор на то, чтобы составлять разнообразные планы - графики, закупки, риски, подготовительные работы.
21	Информационная система клининговой компании	MongoDB	Синтетические данные	Задача - сделать сервис для бизнеса, который занимается уборкой.
22	База данных заболеваний, анализов, симптомов и синдромов (с возможностью логического вывода)	Neo4j	датасет, но можно и другой источник	Задача - сделать базу данных на базе открытого датасета, где можно ставить диагнозы по набору симптомов, составлять перечень анализов, оценивать риски и тд.
23	Сервис-обертка вокруг google drive (для экспорта и более удобного управления)	MongoDB	Синтетические данные	Задача - сделать сервис, который оборачивает api google drive, чтобы в своей БД зафиксировать все документы, права на них и атрибуты. При этом сервис также должен уметь через api выполнять массовые действия с документами (поиск, изменение прав доступа), а также показывать статистику.
24	Видеохостинг	MongoDB	Синтетические данные	Задача - сделать прототип небольшого видеохостинга, на котором можно размещать короткие ролики.

№	Название	Предлагаемая СУБД	Набор данных/API	Комментарий
25	Агрегатор информации о микроклимате в большом помещении (см пример с занятия по Influx - мониторинг микроклимата в ВУЗе)	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.
26	Информационная система компании по производству и монтажу кухонных гарнитуров	MongoDB	Синтетические данные	Задача - сделать сервис для бизнеса, который занимается уборкой. Основной акцент - прототип системы управления производством (заказы, раскрой, рабочие часы, логистика).
27	Сервис транскрипции видео и аудио	MongoDB	Синтетические данные	Задача - сделать сервис для организации транскрипции видео и аудиоматериалов с помощью некоторой распределенной инфраструктуры (это означает, что ресурсы для транскрипции определяются динамически (администраторами) и система должна уметь планировать обработку входных данных исходя из текущих мощностей). Пользователи - администраторы, пользователи. Необходимые (но не достаточные фичи) - страница отдельного задания, профили пользователей, статистика (для админов, по отдельным серверам транскрипции, по пользователям, по задачам), интерфейсы оплаты на заглушках. Саму транскрипцию делать не обязательно, но если очень хочется - можно использовать whisper.ai

№	Название	Предлагаемая СУБД	Набор данных/API	Комментарий
28	Юридические документы	Neo4j	Синтетические данные	Смысл проекта - предоставить интерфейс для хранения, версионирования и анализа (грубого) юридических документов на русском языке. Варианты анализа: разбиение на предложения, абзацы, идентификация пунктов по ключевым словам, поиск формулировок-обязательств / требованийНеобходимые (но не достаточные) фичи - таблица с поиском по документам, страница документа, статистика по документу, общие термины документов
29	Карта транспортной доступности водоемов севера Ленинградской области	Neo4j	Открытые данные OpenStreetMap	Задача в рамках ИДЗ - сделать систему для построения и анализа графа транспортной доступности водоемов (реки, озера, водохранилища) севера ЛО. Необходимые (но не достаточные) фичи - список водоемов, страница визуализации графа , анализ графа (сортировка по доступности, изолированные узлы).
30	Построение маршрутов с учетом плохой парковки	Neo4j	Открытые данные OpenStreetMap	Необходимо взять в качестве демо данных квадрат карты, где находится Петроградский район и для него создать приложение, которое позволит строить маршруты из точки А в точку Б так, чтобы 1) при подъезде к точке Б нужно вести водителя не наикратчайшим путем, а кругами вдоль перспективных для парковки мест 2) при всем этом поездка по маршруту не была супердолгой. Необходимые (но не достаточные) фичи - карта с возможностью задавать точки и выполнять построение маршрута, таблица для визуализации, страница подробной статистики маршрута

Темы 2025

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
01	Каталог хлеба с рецептами	Простой	MongoDB	Синтетические данные	Задача - создать каталог видов хлеба, рецептов, отзывов, комментариев, обсуждений и рекомендаций.
02	БД для фанатов Поттерианы	Простой	Neo4j	https://potterdb.com/	Задача - сделать сервис для фанатов, где они смогут ознакомиться о разных героях / объектах / событиях из вселенной Гарри Поттера, а также визуализировать их взаимосвязи в виде графа.
03	ИТ-каталог	Простой	Neo4j	Gtithub?	Задача - сделать сервис, где будет база данных по всевозможным ИТ профессиям, навыкам, инструментам, технологиям. Интересно построить граф, в котором можно будет увидеть как взаимосвязи разных понятий, так и их востребованность на текущий момент (например, по данным github или открытых рейтингов).

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
04	Каталог фантастики	Простой	MongoDB	IMDB	Задача - сделать сервис, где будет собрана информация про книги, фильмы и сериалы в жанре фантастики и фентези. Нужно построить базу, где все произведения будут тщательно разбиты по жанрам, темам, тегам, взаимосвязям. Организовать интерфейсы для рекомендации произведений по вкусам пользователя.
05	Сервис для огородников	Простой	MongoDB		Задача - сделать сервис, в котором продвинутый огородник может вести записи (какие у него грядки, когда и что сажал, как ухаживал, как это выглядит на фото) и получать полезные рекомендации (когда что делать на участке).

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
06	База данных остатков столярного производства	Простой	Neo4j	Синтетические данные	Задача - создать сервис, где столяры смогут размещать информацию о том, какие у них есть остатки с производства (обрезки, некондиция, брак, опилки и прочее такое - с фото, габаритами и параметрами), примерные цены, адреса, где забрать. Нужно поддержать механизм отзывов, комментариев, страницы как покупателей, так и мастеров.
07	База данных стоматологических карточек	Простой	MongoDB	Синтетические данные	Задача - сделать сервис карточек для стоматологической клиники.
08	Информационная система ветеринарной клиники	Простой	Neo4j	Синтетические данные	Задача - создать ИС для сетевой ветеринарной клиники (есть много филиалов / докторов / услуг, принимаем разных животных, разные процедуры / лекарства). Нужно реализовать юзкейсы для клиентов / докторов / администратора.
09	Скупка электроники	Простой	MongoDB	Синтетические данные	Задача - создать сервис для сервиса скупки электроники. Роли пользователей - администратор (он же скупщик) и клиент. Клиенты отправляют заявки на оценку их электроники, скупщики оценивают по фото или планируют выезд к клиенту.

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
10	Информационная система центра полиграфических услуг	Простой	Neo4j	Синтетические данные	Задача - сделать сервис для обеспечения работы центра полиграфии (печать, сканирование, ризография). Роли: клиенты, сотрудники, администраторы.
11	Хаб для ml датасетов	Простой	MongoDB	Kaggle	Задача - разработать информационную систему для хранения и визуализации набора датасетов для ML.
12	Сервис для текстовых фрилансеров	Простой	MongoDB	Синтетические данные	Задача - сделать сервис, где Заказчики могут найти Фрилансеров (редакторы, писатели). Заказчики создают заказы, оценивают их выполнение, расплачиваются. Фрилансеры берутся за заказы, сдают их по этапу, оценивают Заказчиков.
13	Сервис для организации субботников	Сложный	ArangoDb	Синтетические данные	Задача - сделать сервис, где можно как самому организовать субботник, так и присоединится к чужому субботнику. В сервисе должны быть помимо стандартных страниц (Профили пользователей, поиск, комментарии, отзывы) также и страницы с персональной и общей статистикой (кто и где участвовал, сколько кв км убрали ...).

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
14	Информационная система оконного завода	Сложный	ArangoDB	Синтетические данные	Задача - создать информационную систему для работы оконного завода. Роли пользователей - замерщики, администраторы, покупатели. Замерщики выезжают на адрес, измеряют параметры окон и создают заказы. Администраторы наблюдают за процессом изготовления окон и управляют доставкой до клиента / коммуникацией с ним.
15	Сервис билетов и контролеров общественного транспорта	Сложный	Neo4j	Данные правительства Санкт-Петербурга	Задача - сделать информационную систему для оплаты общественного транспорта в Санкт-Петербурге. Нужны юзкейсы как для пассажиров (кто когда и за что платил, чем расплачивался, балансы, счета), так и для контролеров (у кого когда смена, кого и когда поймал, какой выписал штраф, какие штрафы оплачены, какие нет)

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
16	Каталог схем для вязания и вышивки с аналитикой	Сложный	MongoDB	Схемы из интернета	Задача - реализовать сервис, где можно создавать, хранить и редактировать схемы для вязания / вышивки, а также проводить их анализ на предмет объема необходимого материала, сложности, параметров изделия.
17	Сервис платных парковок	Сложный	MongoDB	Данные правительства Санкт-Петербурга	Задача - сделать информационную систему для оплаты платных парковок в Санкт-Петербурге. Зоны парковки, паркоматы, оплаты, аккаунты, контролеры.
18	Графовый трекер задач	Сложный	Neo4j	Синтетические данные	Задача - реализовать трекер задач (по аналогии с asana trello jira), который бы учитывал и активно использовал зависимости между задачами, а также нелинейность множества задач (например, задачи на стройке). Хранение, визуализация задач в виде таблицы и графа. Анализ графовых характеристик задач, построение списка приоритетов выполнения набора задач.

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
19	Информационная система оценок и аналитика проблемных студентов	Сложный	MongoDB	Учебные программы ФКТИ	Задача - организовать сервис, где будут агрегироваться все действия студента в ВУЗе (учебные активности, посещаемость ВУЗа, участие в социальной активности), и на их основе будут строиться прогнозы по успеваемости, рисках для студента / группы / потока.
20	Информационная система для пожарных бригад - Вызовы, Бригады, Автомобили, Оборудование, Оценка ущерба, Звонки на горячую линию, Сотрудники (Операторы горячей линии, Пожарные)	Сложный	Neo4j	Синтетические данные + Открытые данные OpenStreetMap	Задача - создать сервис, где будет храниться вся необходимая информация для работы одного пожарного депо, а также необходимые веб-страницы, чтобы с этими данными работать.
21	Анализатор студ-отчётов по упоминанию слов и фраз	Сложный	Neo4j	Дипломы МОЭВМ	Задача - сделать сервис, который принимает на вход дипломные работы (docx), парсит их содержимое и разбивает текст по разделам / подразделам, считает статистику (когда и какие слова употребляются, где много шумовых слов и тд), анализирует тексты на схожесть, водность, степень раскрытия отдельных тем / задач.

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
22	Хранилище исходного кода с учетом AST деревьев	Сложный	ArangoDB	Github	Задача - создать сервис, который хостит гит репозитории с исходным кодом И обрабатывает код для конвертации его в AST дерева, чтобы делать анализ и прогнозировать будущие проблемы у разработчиков.
23	Сервис-обертка вокруг google drive (для экспорта и более удобного управления)	Сложный	MongoDB	Синтетические данные	Задача - сделать сервис, который оборачивает api google drive, чтобы в своей БД зафиксировать все документы, права на них и атрибуты. При этом сервис также должен уметь через api выполнять массовые действия с документами (поиск, изменение прав доступа), а также показывать статистику.
24	Редактор для проектирования инженерных коммуникаций в домах	Сложный	Neo4j	Синтетические данные	Задача - реализовать веб сервис для хранения, редактирования и анализа проектов инженерных коммуникаций для загородных домов. Виды коммуникаций - электрика, водопровод, канализация, отопление, вентиляция.
25	Видеохостинг	Сложный	MongoDB	Синтетические данные	Задача - сделать прототип небольшого видеохостинга, на котором можно размещать короткие ролики.

№	Название	Сложность	Предлагаемая СУБД	Набор данных/API	Комментарий
26	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.
27	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).
28	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.info/> - **МОЭВМ Вики** [se.moevm.info]

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

Last update:

