Оленчикова Т.Ю.

Семинар по курсу **«Администрирование и проектирование хранилищ больших данных»**

## Обсуждаемые вопросы

Основные проблемы современных хранилищ данных, решения:

- неудовлетворительная обработка «грязных» данных;

- неудовлетворительные производительность и масштабируемость

- неудовлетворительный выбор источников данных для загрузки в хранилище данных

## Примерные темы докладов

1. Проблемы классических реляционных СУБД. Решения NoSQL
2. Требования к современныч хранилищам данных (ХД). Типы ХД
3. Отечественные облачные ХД.
4. Резидентные базы данных. Где применяются.
5. СУБД для хранилищ данных. Современные решения
6. СУБД PostgreSQL
7. СУБД Riak
8. СУБД Hbase
9. СУБД MongoDB
10. СУБД CouchDB
11. СУБД Neo4j
12. СУБД Redis

## Методические указания

Семинарские занятия строятся на системе докладов (сообщений), которые готовятся студентами по заранее выбранной ими теме. Примерные темы докладов приведены ниже. В то же время студент может, по согласованию с преподавателем, самостоятельно сформулировать тему для своего выступления. При подготовке доклада (сообщения) к семинарскому занятию целью является проведение сравнительного анализа проблемных ситуаций, а также способы и специфика решения перечисленных проблем, где основная задача для студента – самостоятельный анализ.

Перечень требований к выступлению студента на семинаре: связь выступления с предшествующей темой или вопросом; раскрытие сущности проблемы; методологическое значение для научной, профессиональной и практической деятельности. Важнейшие требования к выступлениям студентов - самостоятельность в подборе фактического материала и аналитическом отношении к нему, умение рассматривать примеры и факты во взаимосвязи и взаимообусловленности, отбирать наиболее существенные из них. Приводимые участником семинара примеры и факты должны быть существенными, по возможности перекликаться с профилем обучения.

Выступление студента должно соответствовать требованиям логики. Четкое вычленение излагаемой проблемы, ее точная формулировка, неукоснительная последовательность аргументации именно данной проблемы, без неоправданных отступлений от нее в процессе обоснования, безусловная доказательность, непротиворечивость и полнота аргументации, правильное и содержательное использование понятий и терминов. При оценке докладов используются наиболее существенные критерии: глубина и самостоятельность проведенного студентом анализа проблемных ситуаций отечественной и зарубежной управленческой практики.

В то же время, доклад должен иметь четкую структуру. Всякое выступление на семинаре должно начинаться с изложения его плана. Изложение выводов, полученных студентом в результате анализа, должно соответствовать требованиям логики, быть последовательным и доказательным, аргументация – полной и непротиворечивой. Выступление должно исключать неоправданные отступления от темы. Завершением выступления студентов на семинаре являются ответы на вопросы, которые группа готовит по предложенным учебно-методическим комплексом вопросам семинара. Вопросно-ответный прием публичных выступлений входит в критерий оценки выступлений студентов.

Важным критерием при оценке выступления студента на семинарском занятии является соблюдение регламента. Продолжительность выступления не может превышать 10 минут.

Необходимо также помнить о грамотности, стилистической выдержанности речи выступающего, корректности использования специальных терминов. Использование бытовой лексики и слов-паразитов, тавтология, стилистические и грамматические ошибки, а также чтение текста доклада с листа снижают впечатление от выступления и не могут не сказаться на его оценке.

Доклад на семинарском занятии должен включать сопровождение компьютерной презентацией. Содержание презентации должно включать название темы доклада, его план, реферативное изложение его содержания, основные выводы. Презентация может включать демонстрацию схем, таблиц и др. слайдов, служащих для иллюстрации тех или иных положений доклада.

## Критерии оценивания

Оценка складывается из следующих оценок:

1) раскрытие темы доклада

2) четкость изложения,

3) соблюдение регламента

3) ответы на вопросы

Полное раскрытие темы доклада, выводы логичны – 2 балла; тема раскрыта не полностью или отсутствуют выводы – 1 балл; тема не расрыта – 0 баллов.

Четкость и последовательность изложения – 1 балл, иначе – 0 баллов

Соблюдение регламента – 1 балл, иначе – 0 баллов

Содержательные ответы на вопросы слушателей – 2 балла: неуверенные ответы – 1 балл; неудовлетворительные ответы – 0 баллов

Итого: максимум 6 баллов

## Контрольные вопросы

1. Что входит в состав хранилища данных?
2. Каковы основные характеристики хранилища данных?
3. В чем разница между базой данных и хранилищем данных?
4. В чем разница между ETL и ELT?
5. В каком виде представляется информация в хранилищах данных?
6. основные механизмы доступа к БД в сети Интернет.
7. Системы оперативной обработки информации – OLTP. Архитектура и назначение.
8. Системы консолидации и аналитической обработки информации – ELT. Архитектура и назначение.
9. Концепция многомерного представления данных – гиперкубы.
10. Измерения и факты в гиперкубах. Правил выбора измерений и фактов.
11. Формализация многомерного представления данных: метки, иерархии, ячейки, меры.
12. Операции над данными в гиперкубах: вращение, сечение (срез), свертка и детализация.
13. Агрегация в гиперкубах: виды агрегации.
14. Агрегация в гиперкубах – оценка числа агрегатов для двумерного случая.
15. Классификация архитектуры хранилищ данных.
16. Многомерные хранилища данных.
17. Различие концепций ХД и особенности построения.

## Литература

1. Виды баз данных. Большой обзор типов СУБД <https://habr.com/ru/companies/amvera/articles/754702/>
2. Эрик, Р. Семь баз данных за семь недель. Введение в современные базы данных и идеологию NoSQL / Р. Эрик, Р. У. Джим. ; под редакцией Ж. Картер ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2013. — 384 с. — ISBN 978-5-94074-866-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/58690 (дата обращения: 14.09.2023). — Режим доступа: для авториз. пользователей.