

WEB-ТЕХНОЛОГИИ - ЛЕКЦИИ

02.03.03 -Математическое обеспечение и администрирование информационных систем,
направленность (профиль) -разработка и администрирование информационных систем

<http://vikchas.ru>

Лекция 7. Тема «WEB-технологии и большие данные»

Часовских Виктор Петрович
доктор технических наук, профессор кафедры
ШИиКМ, ФГБОУ ВО «Уральский государственный
экономический университет

Екатеринбург 2025

История появления и развития концепции больших данных

Предпосылки (1960-1990-е)

- **1950-1960-е годы:** Развитие спутниковых систем СССР, фото и видео из космоса определило первые базы данных и систем управления информацией в интересах обороны страны.
- **1960-1970-е годы:** Создание в других странах баз данных и систем управления информацией
- **1980-е годы:** Появление реляционных баз данных и структурированного языка запросов SQL
- **1990-е годы:** Распространение интернета и начало экспоненциального роста цифровой информации

Зарождение термина “Big Data” (2000-е)

- **2001 год:** Аналитик Doug Laney из META Group (позднее Gartner) описывает проблему роста данных через модель “3V”:
 - Volume (объем)
 - Velocity (скорость)
 - Variety (разнообразие)
- **2005 год:** Появление Hadoop – первой распространенной открытой платформы для обработки больших данных, основанной на технологии MapReduce от Google

Эпоха становления (2010-2015)

- **2010 год:** Термин “большие данные” входит в широкое употребление
- **2011-2012 годы:** McKinsey, IBM, Oracle и другие крупные компании начинают активно использовать термин в маркетинговых целях
- **2013-2014 годы:** Формирование экосистемы технологий вокруг Hadoop (Spark, HBase, Hive)
- **2015 год:** К модели “3V” добавляются новые характеристики:
 - Veracity (достоверность)
 - Value (ценность)

Современный этап (2016-настоящее время)

- **2016-2018 годы:** Интеграция технологий больших данных с облачными сервисами
- **2019-настоящее время:** Демократизация технологий больших данных, развитие средств анализа и визуализации, доступных широкому кругу пользователей

-

Эволюция WEB-технологий в контексте больших данных

Web 1.0 и начальный этап (1990-2004)

- Статические веб-сайты с ограниченным взаимодействием
- Данные генерировались преимущественно компаниями
- Простые технологии хранения и обработки информации

Web 2.0 и рост пользовательского контента (2004-2010)

- Появление социальных сетей и платформ для создания пользовательского контента
- AJAX технологии для асинхронного обмена данными
- Рост объемов неструктурированной информации (тексты, изображения, видео)
- Появление первых инструментов аналитики веб-данных

Web 3.0 и интеграция с большими данными (2010-настоящее время)

- **Облачные технологии:**
 - Amazon Web Services (S3, EMR)
 - Google Cloud (BigQuery, Dataflow)
 - Microsoft Azure (HDInsight, Data Lake)
- **Фреймворки для обработки и анализа данных:**
 - Apache Spark, Flink, Storm для потоковой обработки
 - TensorFlow, PyTorch для машинного обучения
 - Elasticsearch, Solr для полнотекстового поиска
- **Технологии визуализации данных:**

- D3.js, Highcharts для интерактивной визуализации
- Tableau, Power BI для бизнес-аналитики
- Grafana, Kibana для мониторинга и анализа логов
- **API и веб-сервисы для работы с данными:**
 - REST API для обмена данными между системами
 - GraphQL для гибких запросов к данным
 - WebSockets для потоковой передачи данных
 -

Современные тенденции взаимодействия WEB и больших данных

1. Распределенные архитектуры

- Микросервисы для масштабируемой обработки данных
- Serverless-архитектура (AWS Lambda, Azure Functions)
- Edge computing для обработки данных ближе к источнику

2. Реальновременная аналитика

- Stream-processing платформы интегрированные с веб-приложениями
- Progressive Web Apps с возможностями работы с большими наборами данных
- Интерактивные дашборды с обновлением в реальном времени

3. Персонализация и машинное обучение

- Рекомендательные системы на основе пользовательского поведения
- Персонализированный контент и A/B тестирование
- Интеграция ML-моделей в веб-приложения через WebAssembly и TensorFlow.js

4. Приватность и этика данных

- Развитие технологий федеративного обучения (Federated Learning)
- Differential Privacy для защиты конфиденциальных данных
- Технологии согласия на использование данных (GDPR, CCPA)

Будущее взаимодействия WEB-технологий и больших данных

- Дальнейшая интеграция инструментов больших данных с веб-технологиями
- Развитие семантического веба и связанных данных (Linked Data)
- Демократизация инструментов анализа данных через веб-интерфейсы
- Развитие децентрализованных технологий хранения и обработки данных (блокчейн, IPFS)
- Усиление роли ИИ в автоматической обработке и интерпретации веб-данных

Взаимное развитие больших данных и веб-технологий продолжает формировать новую цифровую реальность, создавая как новые возможности, так и вызовы для общества, бизнеса и отдельных пользователей.

WEB-технологии в экопространстве больших данных

В современном мире объемы информации растут экспоненциально, формируя так называемое “экопространство больших данных” - сложную экосистему, где данные генерируются, обрабатываются, хранятся и используются. WEB-технологии играют ключевую роль в этой экосистеме, обеспечивая инфраструктуру для работы с большими данными.

Ключевые WEB-технологии в работе с большими данными

1. Облачные платформы

- **Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform** предоставляют масштабируемую инфраструктуру для хранения и обработки данных
- Сервисы типа **AWS Lambda** и **Azure Functions** позволяют создавать бессерверные приложения для обработки данных

2. API и микросервисная архитектура

- REST и GraphQL API обеспечивают стандартизированный обмен данными
- Микросервисная архитектура позволяет создавать масштабируемые системы для работы с большими объемами данных

3. Фреймворки для визуализации данных

- **D3.js, Tableau, Power BI** позволяют создавать интерактивные веб-визуализации больших данных
- **Plotly, Chart.js** обеспечивают построение графиков в веб-интерфейсах

4. Технологии реального времени

- **WebSockets и Server-Sent Events** для передачи потоковых данных в реальном времени
- **Kafka, RabbitMQ** для организации очередей сообщений

5. JavaScript-фреймворки

- **React, Vue.js, Angular** для создания интерактивных пользовательских интерфейсов
- **Node.js** для backend-обработки данных

Актуальные тренды

1. **Progressive Web Applications (PWA)** для работы с данными в оффлайн-режиме
2. **WebAssembly** для высокопроизводительной обработки данных в браузере
3. **Edge Computing**- перемещение обработки данных ближе к источнику
4. **Federated Learning**- распределенное машинное обучение с сохранением конфиденциальности данных
5. **Семантический веб** связанные данные для улучшения интеграции разнородных данных

Проблемы и вызовы

- Обеспечение безопасности и конфиденциальности данных
- Масштабирование систем для работы с постоянно растущими объемами информации
- Оптимизация производительности при передаче и обработке больших объемов данных
- Обеспечение совместимости различных систем и форматов данных

Развитие WEB-технологий в контексте больших данных продолжает трансформировать способы, которыми мы собираем, анализируем и используем информацию, создавая новые возможности и вызовы для разработчиков и пользователей.