



ИНФОРМАЦИОННЫЕ РЕСУРСЫ РОССИИ

02 [197] 2024

Д. Холкин

Многоголосье молчания **6**

**А. Мельников, А. Бачурин,
Е. Гниломёдов, А. Распопов, К. Дворецкая**
Технологии больших данных
в цифровизации ТЭК **30**

В. Цветкова, В. Кулев
Обеспечение сохранности и доступности
электронных фондов библиотеки **38**





ПРАВИТЕЛЬСТВО
САХАЛИНСКОЙ
ОБЛАСТИ



КВАДРАТ
РЕСУРС



ДАЛЬНЕВОСТОЧНЫЙ
ЭНЕРГЕТИЧЕСКИЙ
ФОРУМ

НЕФТЬ И ГАЗ САХАЛИНА

2024

2–4 октября,
г. Южно-Сахалинск

ВОСТОЧНЫЙ ФОРПОСТ ЭНЕРГЕТИЧЕСКОЙ ОТРАСЛИ: ТЕХНОЛОГИЧЕСКИЕ ПРИОРИТЕТЫ И НОВЫЕ ЦЕЛИ



СЕРВИСЫ И ТЕХНОЛОГИИ
НА ШЕЛЬФЕ



ПЕРЕРАБОТКА
УГЛЕВОДОРОДОВ



ЧИСТОЕ БУДУЩЕЕ.
УСТОЙЧИВОЕ
РАЗВИТИЕ И КЛИМАТ



ЛОГИСТИКА И
МЕЖДУНАРОДНОЕ
СОТРУДНИЧЕСТВО



ИНТЕГРАЦИЯ НАУКИ И
ОБРАЗОВАНИЯ В БИЗНЕС.
ОТ ИССЛЕДОВАНИЙ
К ИННОВАЦИЯМ



ВЫСТАВОЧНАЯ
ПРОГРАММА



МОЛОДЕЖНЫЙ
ФОРУМ



КРЕАТИВНЫЙ
ПОТЕНЦИАЛ
ТЕРРИТОРИЙ



СПЕЦИАЛЬНЫЕ
МЕРОПРИЯТИЯ



WWW.SAKHALINOILGAS.RU
+7 499 350 25 35

ПРИГЛАШАЕМ ПРИНЯТЬ УЧАСТИЕ

Содержание

От редакции

- 5 А. Горшкова**
Большие дискуссии о праве
и сущности искусственного
интеллекта

Искусственный интеллект

- 6 Д. Холкин**
Многоголосье молчания
- 24 Д. Телекова, В. Павлова,
А. Поначугин**
Справедливость и предвзятость
к искусственному интеллекту:
изучение подходов к смягчению
предубеждений и обеспечение
справедливости в процессах
принятия его решений

ТЭК

- 30 А. Мельников, А. Бачурин,
Е. Гниломёдов, А. Распопов,
К. Дворецкая**
Технологии больших данных
в цифровизации ТЭК

Библиотека

- 38 В. Цветкова, В. Кулев**
Обеспечение сохранности
и доступности электронных
фондов библиотеки



ИНФОРМАЦИОННЫЕ
РЕСУРСЫ РОССИИ

Интернет

- 54 Н. Редькина**
Веб-архив: важнейший источник
информации vs неупорядоченное
хранилище веб-контента

Строительство

- 67 А. Пиляй, Е. Гинзбург**
Применение глубокого обучения
для обнаружения повреждений
на фасаде зданий

Цифра

- 75 М. Муконина,
Д. Глазунов, А. Гальцева**
Разработка программного продукта
для стандартизации документов
в электронном виде

- 85 С. Малахов, Д. Якупов**
Анализ производительности
программного протокола OpenFlow
на виртуальной машине

Безопасность

- 93 А. Яковишин, И. Кузнецов,
И. Дроздов, Д. Письменский**
Перспективы развития
информационной безопасности:
глобальные вызовы
и стратегии защиты



Мельников Андрей
Заместитель директора центра
инновационных программ, НИОКР
и отраслевой стандартизации, к. т. н.,
НИИТТ (ООО «НИИ Транснефть»)
E-mail: MelnikovAV@niitnn.transneft.ru

Бачурин Александр
Ведущий научный сотрудник сектора
инновационных программ, к. т. н.,
НИИТТ (ООО «НИИ Транснефть»)
E-mail: BachurinAI@niitnn.transneft.ru

Гниломёдов Евгений
Старший научный сотрудник сектора
инновационных программ, к. э. н.,
НИИТТ (ООО «НИИ Транснефть»)
E-mail: GnilomedovEV@niitnn.
transneft.ru

Распопов Андрей
Заместитель директора центра
инновационных программ, НИОКР
и отраслевой стандартизации,
к. т. н., НИИТТ (ООО «НИИ
Транснефть»)
E-mail: RaspopovAA@niitnn.transneft.ru

Дворецкая Кристина
Выпускник МГИМО
МИД России

ТЕХНОЛОГИИ БОЛЬШИХ ДАННЫХ В ЦИФРОВИЗАЦИИ ТЭК

Аннотация. Развитие технологий больших данных за последние несколько лет стало одним из главных достижений в ИТ-индустрии. Благодаря стремительному развитию информационных технологий в последние годы облачные вычисления и технологии больших данных внедрены и в нефтегазовую отрасль. Добыча и транспортировка углеводородов характеризуется такими особенностями, как большой объем данных, высокая производительность получения и анализа данных, высокая надежность и постоянное расширение массивов хранения данных. Применение облачной платформы нового поколения поможет решить задачи сбора и функционирования баз данных. Облачные вычисления объединяют прикладную программную платформу больших данных. В статье проанализированы методы анализа больших данных, архитектура и применение промышленных больших данных в нефтегазовой отрасли, а также сформулированы выводы об эффективности применения больших данных для оптимизации традиционных технологических операций в нефтегазовой промышленности.

Ключевые слова:

промышленные большие данные, облачные вычисления и технологии, цифровизация, нефтегазовая промышленность.

**Облачные
вычисления
активно
применяются
для оптимизации
традиционных
технологических
операций:
в разведке,
бурении, добыче
и переработке
нефти и газа**

Большие данные представляют собой структурированные и неструктурированные данные огромных объемов и типов представления, эффективная обработка которых требует специализированных методик и программных средств. С появлением новых технологических возможностей анализа огромных массивов данных появилось данное понятие, формализованное в серии научных публикаций конца 2000-х гг., первой из которых считается статья Клиффорда Линча [1].

Бесспорно, огромные массивы информации требуют новейших технологий, позволяющих собирать, анализировать, визуализировать и хранить сложные наборы данных. Во многих научных статьях отмечается, что большие данные сочетают в себе всевозможные виды ИТ-технологий: оперативный поиск, систематический сбор и обработку, статистический анализ, а также вычисления больших массивов данных. Благодаря передовым технологиям, большие данные позволяют определить ранее не исследуемые законы и неизвестные корреляции между разрозненной информацией.

Основные сценарии применения больших данных включают пакетную, потоковую и поисковую обработку. Пакетная обработка отличается накоплением больших объемов данных, которые обрабатываются по запросу от пользователей

и преимущественно применяется в интернете вещей¹, облачных вычислениях и интернете. При пакетной обработке данные, обладающие высокой точностью и низкой плотностью, передаются в блочном режиме. Данные сначала накапливаются и анализируются, достигнув определенного объема. Кроме того, поисковая обработка включает в себя анализ потоковых данных и интерактивную обработку данных. Потоковые данные характеризуются бесперебойным анализом ранее не хранимых в памяти потоков данных, которые обрабатываются практически в режиме реального времени и результаты которых формируются мгновенно. Типичное применение потоковой обработки данных заключается в сборе данных – получение журналов, показаний датчиков и прочее в режиме реального времени.

Интерактивная обработка данных преимущественно применяется в системах обработки информации (OLTP, OLAP, Hive, Pig) и в интернете (поисковая система, электронная почта, средства коммуникации в реальном времени и т. д.). Крупнейшими известными системами управления базами дан-

¹ Системы, которые объединяют устройства в компьютерную сеть и позволяют им собирать, анализировать, обрабатывать и передавать данные другим объектам через программное обеспечение, приложения или технические устройства.