

**Технический комитет
по стандартизации
«Искусственный интеллект» (ТК 164)**

Подкомитет «Данные» (ПК 02)

СТАНДАРТЫ В ОБЛАСТИ БОЛЬШИХ ДАННЫХ

Юрий Евгеньевич Хохлов

председатель подкомитета

Что такое большие данные?



Большие массивы данных, отличающиеся главным образом такими характеристиками, как объем, разнообразие, скорость обработки и/или вариативность, которые требуют использования технологии масштабирования для эффективного хранения, обработки, управления и анализа

Источник: Data Never Sleeps 8.0
<https://www.domo.com/>

Источник: ISO/IEC 20546:2019 "Information technology — Big data — Overview and vocabulary"

Работа с большими данными



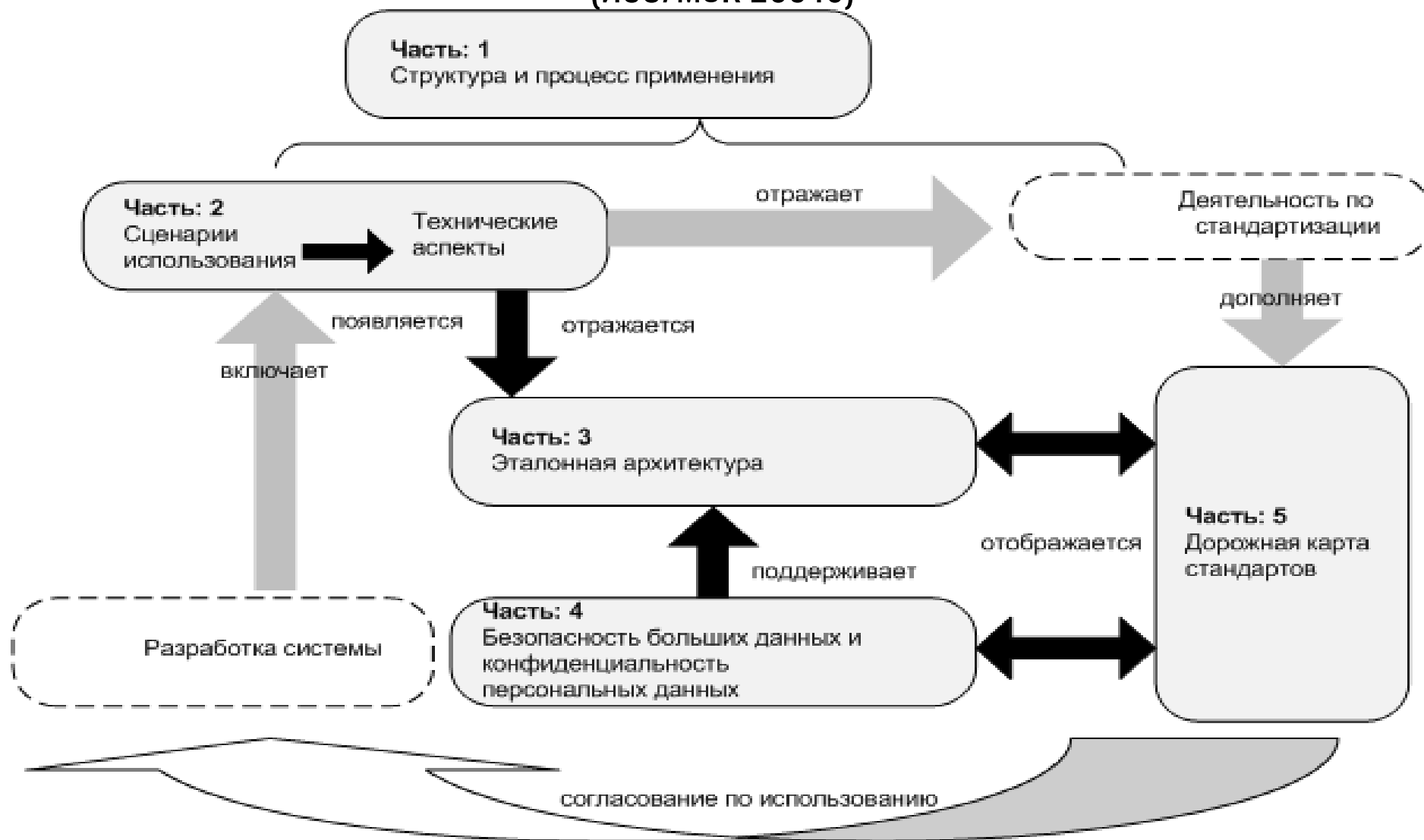
Эталонная архитектура больших данных - Роли



Источник: ISO/IEC 20547-3:2020 "Information technology — Big data reference architecture — Part 3: Reference architecture"

Взаимосвязи между частями стандарта ИСО/МЭК 20547-Х

Терминологическая основа - ГОСТ Р «Информационные технологии. Большие данные. **Обзор и словарь**» (ИСО/МЭК 20546)



Работы над международными стандартами в области больших данных

- ✓ Завершается разработка стандарта ISO/IEC 24668 «Information technology – Artificial intelligence – Process management framework for Big data analytics»
- ✓ Начинается работа над новой серией стандартов по качеству данных для аналитики и машинного обучения:
 - ISO/IEC WD 5259-1 Data quality for analytics and ML – Overview
 - ISO/IEC WD 5259-2 Data quality for analytics and ML – Data quality measures
 - ISO-IEC WD 5259-3 Data quality for analytics and ML – Management
 - ISO-IEC WD 5259-4 Data quality for analytics and ML – Process

Стандартизация больших данных в РФ

Действующие игроки

- Ассоциация интернета вещей
- Российская венчурная компания
- Рабочая группа 3 «Большие данные» ТК 194 «Кибер-физические системы»



Новые игроки

- Ассоциация больших данных
- Ассоциация участников рынка данных
- Центр хранения и анализа больших данных МГУ имени М.В. Ломоносова
- Институт развития информационного общества
- Подкомитет «Данные» ТК 164 «Искусственный интеллект»

Текущая деятельность ПК 02 «Данные»

- ГОСТ Р «Информационные технологии. Большие данные. Обзор и словарь» (ISO IEC 20546:2019 «Information technology - Big data reference architecture -Overview and vocabulary»). **Нормоконтроль**
- ГОСТ Р «Информационные технологии. Эталонная архитектура больших данных. Часть 1. Структура и процесс применения» (ISO/IEC TR 20547-1:2020 «Information technology — Big data reference architecture — Part 1: Framework and application process»). **27 февраля 2021 г. завершение публичного обсуждения**
- ГОСТ-Р «Информационные технологии. Эталонная архитектура больших данных. Часть 2. Варианты использования и производные требования» (ISO/IEC TR 20547-2:2018 «Information technology — Big data reference architecture — Part 2: Use cases and derived requirements»). **Подготовка окончательной редакции**
- ГОСТ Р «Информационные технологии. Эталонная архитектура больших данных. Часть 5. Направления стандартизации» (ISO IEC TR 20547-5:2018 «Information technology - Big data reference architecture - part: 5: Standards roadmap»). **Начало публичного обсуждения**
- ГОСТ Р «Информационные технологии. Большие данные. Техническое задание. Требования к содержанию и оформлению». **Подготовка окончательной редакции**

История создания BD4DE-MM

- ✓ Центр хранения и анализа больших данных создан в 2018 г.
- ✓ «Мониторинг и стандартизация развития и использования технологий хранения и анализа больших данных в цифровой экономике Российской Федерации» - один из ключевых проектов Центра
- ✓ В 2019 г. разработана модель зрелости работы с большими данными в организации (Big Data for Digital Economy Maturity Model, BD4DE-MM)
- ✓ В 2020 г. на основе BD4DE-MM проведено пилотирование методики (само)оценки уровня зрелости работы с большими данными в организации
- ✓ В 2020 г. разработан онлайн-инструментарий для (само)оценки



О ЦЕНТРЕ

НОВОСТИ

КОНСОРЦИУМ

ПРОЕКТЫ

ЦЕНТР ХРАНЕНИЯ И АНАЛИЗА БОЛЬШИХ ДАННЫХ

Центр компетенций ИТИ по направлению «Технологии хранения и анализа больших данных» на базе МГУ имени М.В. Ломоносова создан в рамках государственной программы поддержки Центров компетенций ИТИ. Оператор проекта – АО «Российская венчурная компания».

<http://bigdata.msu.ru>

BD4DE-MM – Концептуальная схема

5 уровней зрелости: набор свойств, которыми обладает организация

5	Есть стратегия работы с большими данными и план реализации Участствует в разработке нормативного правового и технического регулирования, саморегулирования	Сотрудники обладают навыками работы с большими данными Система повышения уровня компетенций кадров Руководители являются "лидерами" по работе с большими данными	Процессы работы с большими данными регламентированы Руководитель и подразделение, отвечающие за работу с большими данными Проектное управление при работе с большими данными	Вся деятельность организации обеспечена большими данными Совместное использование больших данных для всей деятельности Работа с большими данными на всех стадиях жизненного цикла	Используются специальные инструменты работы с большими данными Регулярно используется аналитика больших данных	Существующая инфраструктура работы с большими данными полностью обеспечивает всю деятельность организации	Достигнуты эффекты от использования больших данных: рост выручки, повышение качества товаров и услуг, снижение экологического ущерба, улучшение работы с кадрами
4	Есть стратегия работы с большими данными и план реализации Участствует в разработке нормативного правового и технического регулирования, саморегулирования	Сотрудники обладают навыками работы с большими данными Система повышения уровня компетенций кадров Руководители являются "лидерами" по данным	Процессы работы с большими данными регламентированы Руководитель и подразделение, отвечающие за работу с большими данными	Вся деятельность организации обеспечена большими данными Совместное использование больших данных для всей деятельности	Используются специальные инструменты работы с большими данными Регулярно используется аналитика больших данных	Существующая инфраструктура работы с большими данными полностью обеспечивает всю деятельность организации	Достигнуты эффекты от использования больших данных: рост выручки, повышение качества товаров и услуг, снижение экологического ущерба, улучшение работы с кадрами
3	Есть стратегия работы с большими данными и план реализации Участствует в разработке нормативного правового и технического регулирования, саморегулирования	Сотрудники обладают навыками работы с большими данными Система повышения уровня компетенций кадров Руководители являются "лидерами" по работе с большими данными	Процессы работы с большими данными регламентированы Руководитель и подразделение, отвечающие за работу с большими данными Проектное управление при работе с большими данными	Вся деятельность организации обеспечена большими данными Совместное использование больших данных для всей деятельности Работа с большими данными на всех стадиях жизненного цикла	Используются специальные инструменты работы с большими данными Регулярно используется аналитика больших данных	Существующая инфраструктура работы с большими данными полностью обеспечивает всю деятельность организации	Достигнуты эффекты от использования больших данных: рост выручки, повышение качества товаров и услуг, снижение экологического ущерба, улучшение работы с кадрами
2	Есть стратегия работы с большими данными и план реализации Участствует в разработке нормативного правового и технического регулирования, саморегулирования	Сотрудники обладают навыками работы с большими данными Система повышения уровня компетенций кадров Руководители являются "лидерами" по работе с большими данными	Процессы работы с большими данными регламентированы Руководитель и подразделение, отвечающие за работу с большими данными Проектное управление при работе с большими данными	Вся деятельность организации обеспечена большими данными Совместное использование больших данных для всей деятельности Работа с большими данными на всех стадиях жизненного цикла	Используются специальные инструменты работы с большими данными Регулярно используется аналитика больших данных	Существующая инфраструктура работы с большими данными полностью обеспечивает всю деятельность организации	Достигнуты эффекты от использования больших данных: рост выручки, повышение качества товаров и услуг, снижение экологического ущерба, улучшение работы с кадрами
1	Есть стратегия работы с большими данными и план реализации Участствует в разработке нормативного правового и технического регулирования, саморегулирования	Сотрудники обладают навыками работы с большими данными Система повышения уровня компетенций кадров Руководители являются "лидерами" по работе с большими данными	Процессы работы с большими данными регламентированы Руководитель и подразделение, отвечающие за работу с большими данными Проектное управление при работе с большими данными	Вся деятельность организации обеспечена большими данными Совместное использование больших данных для всей деятельности Работа с большими данными на всех стадиях жизненного цикла	Используются специальные инструменты работы с большими данными Регулярно используется аналитика больших данных	Существующая инфраструктура работы с большими данными полностью обеспечивает всю деятельность организации	Достигнуты эффекты от использования больших данных: рост выручки, повышение качества товаров и услуг, снижение экологического ущерба, улучшение работы с кадрами

38 аспектов: характеристики, которые можно наблюдать на каждом уровне зрелости работы с большими данными в организации

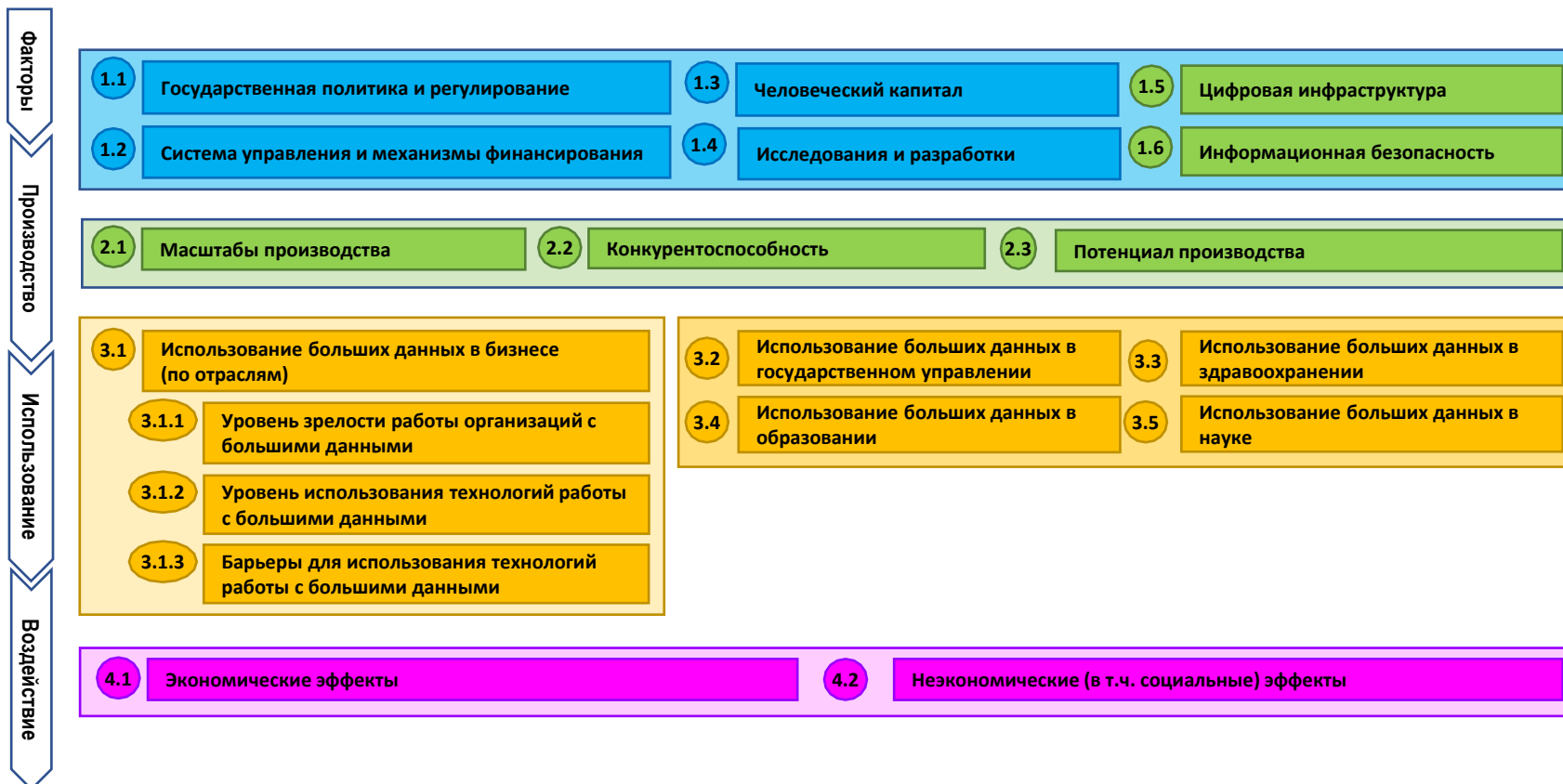
STR Стратегия и регулирование	HR Кадры и лидерство	ORG Организация работы	DATA Данные	BDA Инструменты и аналитика	INFR Инфраструктура и безопасность	IMP Воздействие
---	--------------------------------	----------------------------------	-----------------------	---------------------------------------	--	---------------------------

7 размерностей: логическая группировка аспектов работы с большими данными

Работа с большими данными



Многомерная модель мониторинга развития и использования технологий работы с большими данными



16 направлений, соответствующих 4 компонентам:

- Экономическое и социальное воздействие больших данных
- Использование технологий работы с большими данными
- Производство технологий, решений и услуг для работы с большими данными
- Факторы, влияющие на производство, использование и воздействие технологий работы с большими данными

Индустрия больших данных

Разработчики технологий, решений и услуг для работы с большими данными

Разработка решений и услуг для работы с большими данными

Аудит решений и услуг для работы с большими данными

Системная интеграция для работы с большими данными

Поставщики больших данных

Поставщики решений и услуг по стадиям жизненного цикла работы с большими данными

Комплектование массивов большими данными

Предобработка больших данных

Аналитика больших данных

Визуализация больших данных

Доступ к большим данным

Поставщики решений и услуг для интеграции и взаимодействия

Поставщики решений и услуг для управления большими данными

Поставщики решений и услуг для обеспечения безопасности БД

Поставщики комплексных решений и услуг для работы с большими данными

Поставщики инфраструктурных решений и услуг для работы с большими данными

Вычислительные ресурсы для работы с большими данными

Платформы для работы с большими данными

Инфраструктура передачи больших данных

Приглашение к сотрудничеству

Центр хранения и анализа больших данных на базе МГУ имени М.В. Ломоносова приглашает к сотрудничеству

Партнеры

- ✓ разработчики технологий работы с большими данными
- ✓ пользователи технологий, решений и услуг для работы с большими данными
- ✓ поставщики больших данных
- ✓ исследователи в области больших данных

Направления

- ✓ исследования и разработки в области технологий работы с большими данными в цифровой экономике
- ✓ мониторинг развития и использования технологий работы с большими данными
- ✓ (само)оценка уровня зрелости работы с большими данными в организации

Хохлов Юрий Евгеньевич

председатель совета директоров

Института развития информационного общества

yuri.hohlov@iis.ru

Подкомитет «Данные» (ПК 02)

Э-почта: bigdata-wg02@digital.msu.ru

Веб-сайт: <https://bigdata.msu.ru/standards>