

ТК 164  
«Искусственный  
интеллект»

Подкомитет 02  
«Данные»

О разработке  
национального  
стандарта ГОСТ Р  
«Информационные  
технологии.  
Большие данные.  
Обзор и словарь»  
Е.В. Иванова

# Обзорная информация

ГОСТ Р «Информационные технологии. Большие данные. Обзор и словарь»

ISO IEC 20546  
«Information technology –  
Big Data reference  
architecture – Overview  
and vocabulary»

- Цель стандарта –  
установление  
стандартизированных  
терминов, обеспечение  
единства терминологии  
и взаимопонимания в  
предметной области  
«Информационные  
технологии – Большие  
данные»

## Разработчики

- МГУ имени  
М.В.Ломоносова
- АНО «Институт  
развития  
информационного  
общества»

# Основания для разработки

создание базовых условий развития индустрии больших данных для:

- нормативного применения в Российской Федерации международных и национальных стандартов в области больших данных;
- повышения конкурентоспособности отечественных продуктов в мире;
- единого понятийного аппарата в области технологий хранения и анализа больших данных;
- повышения эффективности использования технологий работы с большими данными за счёт унификации нормативно-технических требований.

# Этапы реализации

- Январь – апрель 2020 г: перевод стандарта ISO/IEC 20546:2019 "Information technology – Big data – Overview and vocabulary" и его научное редактирование
- 13 мая 2020 г. опубликована первая редакция стандарта на сайте Центра компетенций НТИ по технологиям хранения и анализа больших данных по ссылке
- 13 июля 2020 окончание приема замечаний. Собрано и обработано 54 замечания от экспертов
- 24 августа 2020 г. Окончательная редакция была направлена в Технический комитет 164 для рассылки экспертам
- 15 сентября подготовлена окончательная версия стандарта

# Принципы адаптации. Источники

## Принципы

- Отсутствие русскоязычной нормативной терминологии в области БД и ИИ
- Уровни унификации и нормализации терминологии: логический, содержательный, лингвистический
- Стремление к русификации, отказ от калькированных терминов
- Точное, однозначное соотнесение с понятиями и объектами

## Источники

- ГОСТ 33707-2016 (ИСО/МЭК 2382:2015) Информационные технологии (ИТ). Словарь
- ГОСТ 17788-2016 Информационные технологии (ГОСТ ИСО/МЭК ИТ). Облачные вычисления. Общие положения и терминология
- ГОСТ Р ИСО/МЭК 19784-4-2014 Информационные технологии (ИТ). Биометрия. Биометрический программный интерфейс. Часть 4. Интерфейс поставщика функции биометрического датчика

# Ценность

- Формирование единого понятийного аппарата для научной и прикладной деятельности в области БД
- Формирование понятийной семантической системы научной области
- Создание условия для развития отрасли (понимание – важнейшая составляющая эволюции)

# Результат

- Утверждены и нормализованы русскоязычные понятия, применяемые в области больших данных, и имеющие отношение к их сбору, анализу и обработке
- Сформулированы адаптированные дефиниции для ключевых терминов предметной области «большие данные» (velocity, variety, volume, veracity, variability – скорость обработки, изменчивость, объем, достоверность, вариативность).
- Обеспечена терминологическая основа для стандартов, связанных с большими данными
- Заложена база для формирования и развития русскоязычного терминологического аппарата в области больших данных

## Контакты

Научный редактор  
стандарта

Елизавета Васильевна  
Иванова

[elizaveta@egotranslating.ru](mailto:elizaveta@egotranslating.ru)