

## Аннотация дисциплины М.1.2.4 Дисциплина. Технологии и инструментальные средства анализа больших данных

Дисциплина "Технологии и инструментальные средства анализа больших данных" изучается обучающимися по основной профессиональной образовательной программе "Искусственный интеллект в бизнес-аналитике" направления подготовки "09.04.03 Прикладная информатика".

Дисциплина изучается в 2, 3 семестре. Общая трудоемкость дисциплины составляет 144/4 часов/з.ед. Самостоятельная работа заключается в выполнении работ, указанных в разделе 4.

В ходе изучения дисциплины осуществляется текущий контроль в форме технологии рейтингового контроля в соответствии с технологической карты дисциплины, размещенной на электронном курсе, а также промежуточный контроль в форме балльно-рейтинговый контроль, без контрольной акции.

Целью изучения дисциплины является формирование следующих компетенций:

1. ПК-8 Способен осуществлять руководство по созданию и развитию систем и комплексов обработки данных, в т.ч. больших данных, для корпоративных и государственных заказчиков
2. ПК-9 Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях

В ходе изучения дисциплины последовательно рассматриваются темы:

1. Большие данные: общие понятия, классификация, примеры.  
Источники больших данных: показания датчиков, приборов и других устройств; Интернет (соцсети, форумы, блоги, СМИ и т.д.); корпоративные архивы документов.  
Методы анализа больших данных: Data Mining, краудсорсинг, смешение и интеграция данных, машинное обучение, искусственные нейронные сети, распознавание образов, прогнозная аналитика, имитационное моделирование, пространственный анализ, статистический анализ, визуализация данных.  
Структура центров обработки данных. Действующая нормативная база в области проектирования и строительства центров обработки данных.
2. Технологии анализа больших данных: модель MapReduce, технология Hbase (<https://habr.com/ru/company/dca/blog/280700/>), (<https://habr.com/ru/post/267361/>). NoSQL, Hadoop, R, Python. Особенности технологий для анализа больших данных: звуковых, речевых, медицинских, метеорологических, промышленных данных и систем видеонаблюдения и др.
3. Инструментальные средства анализа больших данных: Google Analytics, Mixpanel, Kissmetrics, Weborama, OWOX BI, Datorama, Supermetrics, Funnel, Improvado, AppsFlyer, Adobe Analytics, Fivetran и др. Обзор и возможности инструментальных средств анализа данных.
4. Принципы и методы хранения данных (в том числе звуковых, речевых, метеорологических, промышленных данных и данных систем видеонаблюдения) на общедоступных платформах для обеспечения потребностей организаций-разработчиков в области искусственного интеллекта.  
Принципы разработки проектной документации центра обработки данных. Системная архитектура серверного оборудования и систем хранения данных, цифровых платформ анализа данных.  
Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных для разных сфер и отраслей.

Основными стратегическими образовательными технологиями являются: дискуссионные,

имитационное моделирование, исследовательские, лекционные занятия, практические и лабораторные занятия, процедуры самообучения.

В рамках указанных технологий применяются тактические образовательные технологии: задания, информационные, классическая лекция, лекция с элементами мозгового штурма, мини-проекты, case-study.