

# Интеллектуальные технологии и кибербезопасность цифрового предприятия

«38.04.05 – бизнес информатика  
направленность интеллектуальное управление цифровым предприятием»

<http://vikchas.ru>

**Тема 1. Интеллектуальные технологии цифрового предприятия**  
**Лекция 2 «Об искусственном интеллекте»**  
»

**Часовских Виктор Петрович**  
д-р техн. наук, профессор кафедры ШИиКМ

ФГБОУ ВО «Уральский государственный экономический  
университет»

Екатеринбург 2022

# Области искусственного интеллекта

Искусственный интеллект включает в себя множество областей математики и информационных технологий, а также биологии, физики и других наук.

Мы описали только самые известные и значимые события и подходы, разумеется, идей создания искусственного интеллекта было гораздо больше.

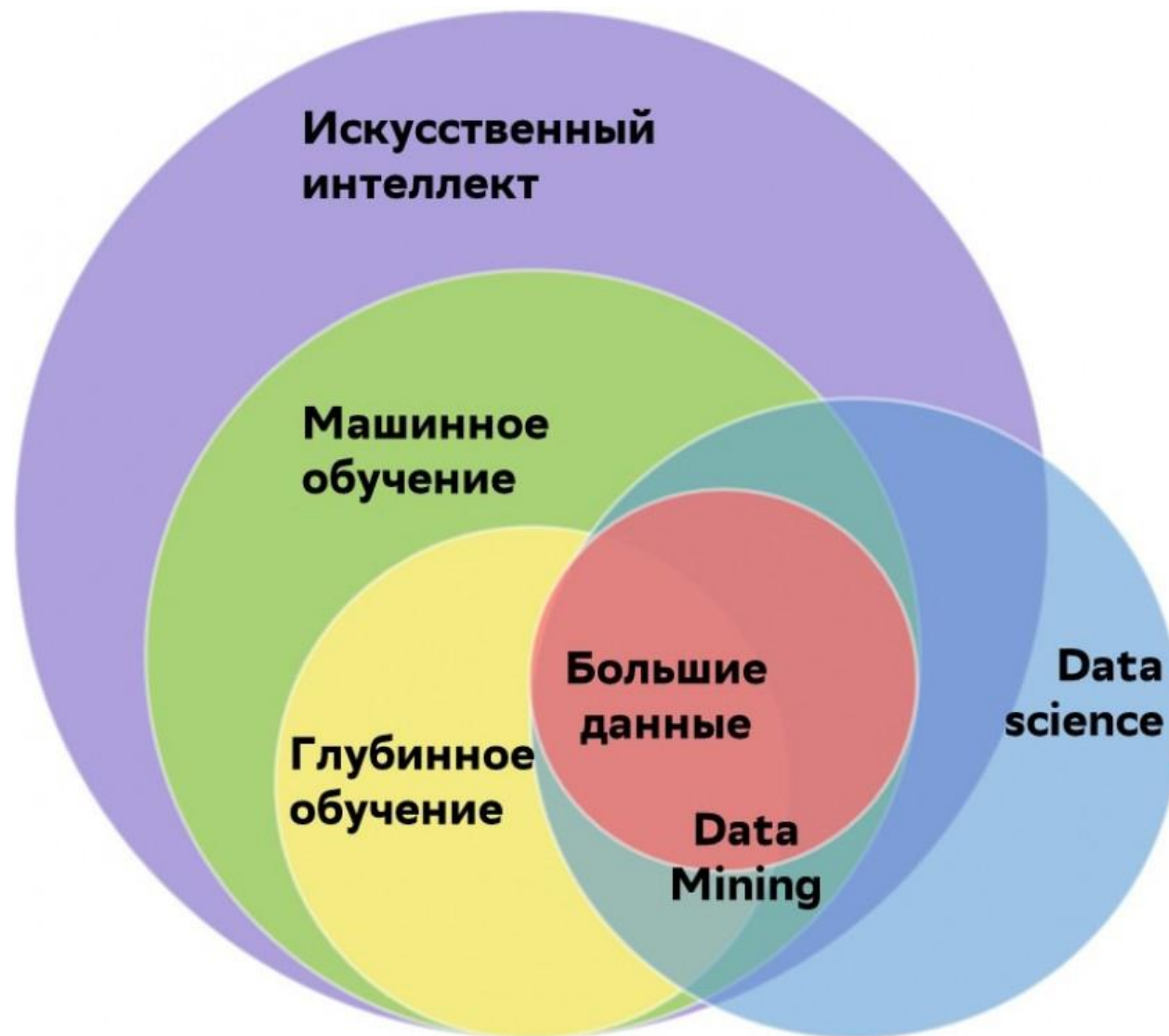
Например, существует подход на основе эволюционных алгоритмов: он заключается в том, чтобы попытаться имитировать «эволюцию» с помощью случайных «мутаций» программы; сегодня такие подходы используются совместно с современными системами искусственного интеллекта, например с нейронными сетями.

Помимо понимания искусственного интеллекта как способности компьютера решать интеллектуальные задачи подобно человеку, существует понимание ИИ как создания компьютера, имитирующего человеческий мозг. Однако в науке пока нет полного понимания, как работает мозг, поэтому способов его искусственного повторения тоже не существует.

На сегодняшний день технологии искусственного интеллекта и обработки больших объемов данных активно используются в бизнесе, конкретным примерам посвящен следующий блок курса.

При этом выделяют несколько смежных областей, отвечающих за разработку этих технологий.

# Смежные области ИИ



## **Машинное обучение**

Класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач.

## **Глубинное обучение**

Иногда называют «глубокое обучение» (от англ. Deep learning). Подобласть машинного обучения, где в качестве алгоритмов используются нейронные сети.

## **Data Science**

Это концепция объединения статистики, анализа данных, машинного обучения и связанных с ними методов для понимания и анализа реальных явлений.

## **Data Mining**

Широкое понятие, означающее извлечение знаний из данных

## Большие данные

Это набор подходов и методов, разработанных для анализа данных огромных размеров.

Стоит отметить, что не всякая работа с данными относится к искусственному интеллекту (например, аналитик, делающий вывод по графикам, не относится к искусственному интеллекту), и не все алгоритмы искусственного интеллекта разрабатываются с использованием данных (например, экспертные системы, упомянутые выше).

Тем не менее, множество современных интеллектуальных систем основано именно на обучении по данным: машинный перевод, распознавание изображений и речи, прогнозирование поведения клиентов и др.

Во время обучения по данным алгоритм «изучает» большое количество реальных случаев (например, поведения клиентов или переводов текстов) и благодаря этому делает качественные предсказания для новых случаев.

# Технологии работы с большими данными

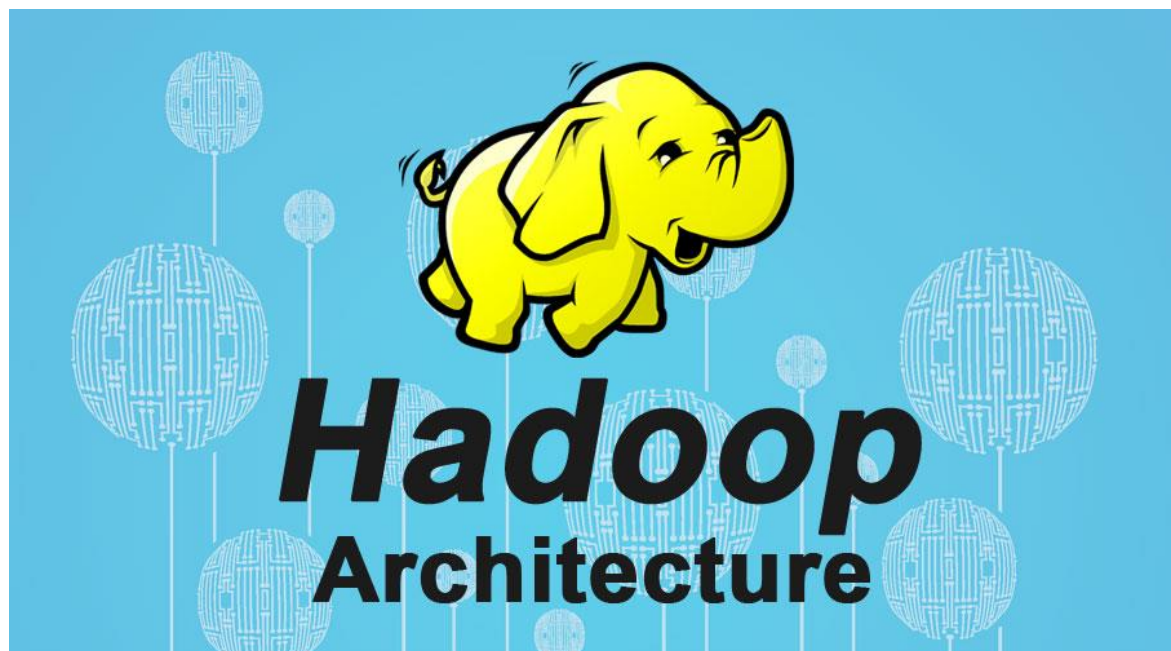
Ранее мы обсудили, что современные технологии искусственного интеллекта практически целиком основываются на работе с данными, и, кроме того, использование данных и аналитика данных сами по себе являются мощным драйвером развития в современном мире. При этом речь идет об огромных массивах данных, обработать которые на персональном компьютере не предоставляется возможным. Об инструментах хранения и обработки больших данных мы говорили в соответствующем курсе, здесь кратко напомним перечень.

Для хранения и обработки больших данных создают специальные дата-центры или арендуют мощности на облачных сервисах, таких как Amazon Web Services или Microsoft Azure.

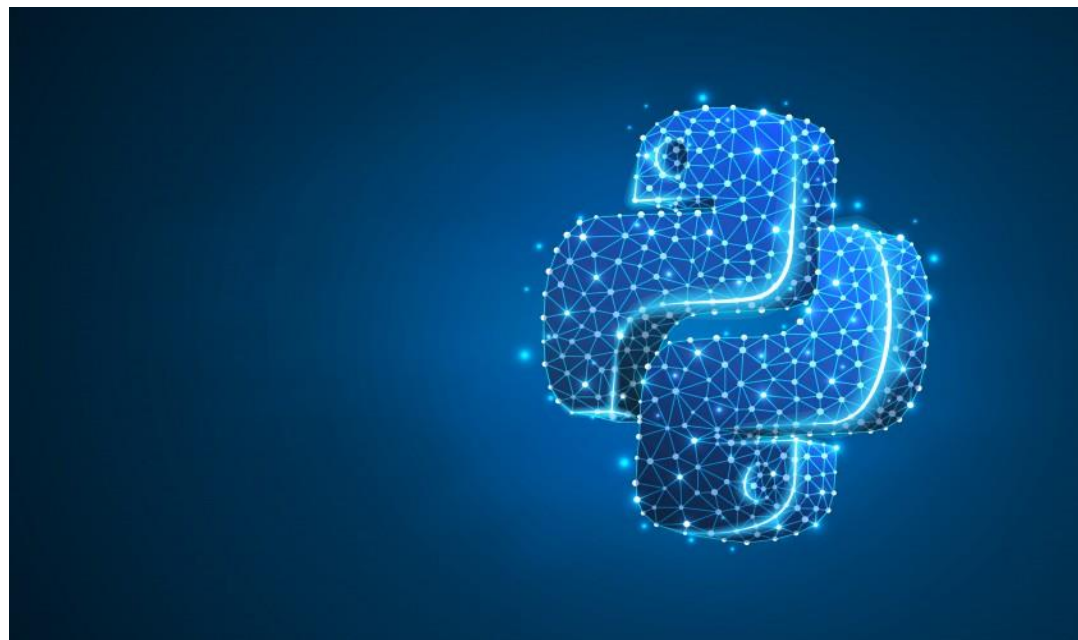


Вместо последовательной обработки данных о миллиардах клиентов используют распределенное хранение данных и параллельную их обработку, например в [Hadoop](#) или [Spark](#).

Отдельный модуль SparkML позволяет строить модели машинного обучения на сверхбольших объемах данных.



Для построения моделей машинного обучения наиболее популярен язык программирования Python и его библиотеки Scikit-learn, LightGBM, CatBoost и другие. Для обучения нейронных сетей используются отдельные библиотеки: PyTorch и TensorFlow. Также существуют графические интерфейсы, например RapidMiner, но они предоставляют более ограниченный функционал по сравнению с Python и поэтому не так популярны.



# Направления применения

Говоря о применении ИИ в бизнесе, сегодня чаще всего имеют в виду применение технологий машинного обучения (machine learning) — именно они позволили достичь высоких результатов в анализе изображений, текстов, в играх и других нетривиальных задачах.

Эти технологии подразумевают извлечение знаний из огромных массивов информации (наборов данных, или по-английски dataset — датасетов).

Принцип работы алгоритма машинного обучения следующий: по большому количеству примеров вида вход — выход настраивают алгоритм, который сможет по входу предсказывать выход. Процесс настройки алгоритма называется обучением (learning).

Разработка способов обучения алгоритмов — это отдельная область прикладной математики на стыке с информационными технологиями.

Однако заказчику досконально понимать, как устроены механизмы машинного обучения, не обязательно, главное — понимать, какие задачи эта технология может решать, и на высоком уровне ориентироваться в области.

Выделяют два основных направления применения ИИ:

**автоматизация**

**и**

**прогнозирование.**

## Автоматизация



ИИ успешно применяется в задачах, в которых вид входных и выходных данных всегда один и тот же, однако четкого алгоритма получения результата не существует. При этом поток входных данных настолько большой, что для решения задачи приходится нанимать большую команду — ИИ же может избавиться от такой необходимости и значительно оптимизировать расходы.

## Примеры автоматизации

Например,

ИИ успешно сортирует отзывы о продукте, категоризирует их, собирает статистику и выделяет негативные отзывы, требующие срочного ответа.

Другой пример — модерация контента на сайтах объявлений: вместо большого отдела модераторов объявления, не соответствующие правилам сервиса, отфильтровывает ИИ.

Аналогичные примеры можно найти в такси или общепите: ИИ следит за скоростью водителя или за качеством мытья рук сотрудниками ресторана. Еще один пример — применение ИИ для организации хранения документов и для их обработки.

В эту же группу можно отнести чат-боты, заменяющие сотрудников контакт-центра, отвечающих на обращения клиентов в контакт-центр.



# Прогнозирование



Речь идет о решении задач, для которых ответ неизвестен, но его можно спрогнозировать на основе исторических данных. Для решения таких задач приходится нанимать опытных специалистов, но человеку нужно много времени, чтобы проанализировать большие объемы информации. ИИ справляется быстрее и качественнее, потому что быстрее выполняет вычисления, «запоминает» больше зависимостей в данных, кроме того, ИИ оценивает ситуацию объективно и беспристрастно.

## Примеры прогнозирования

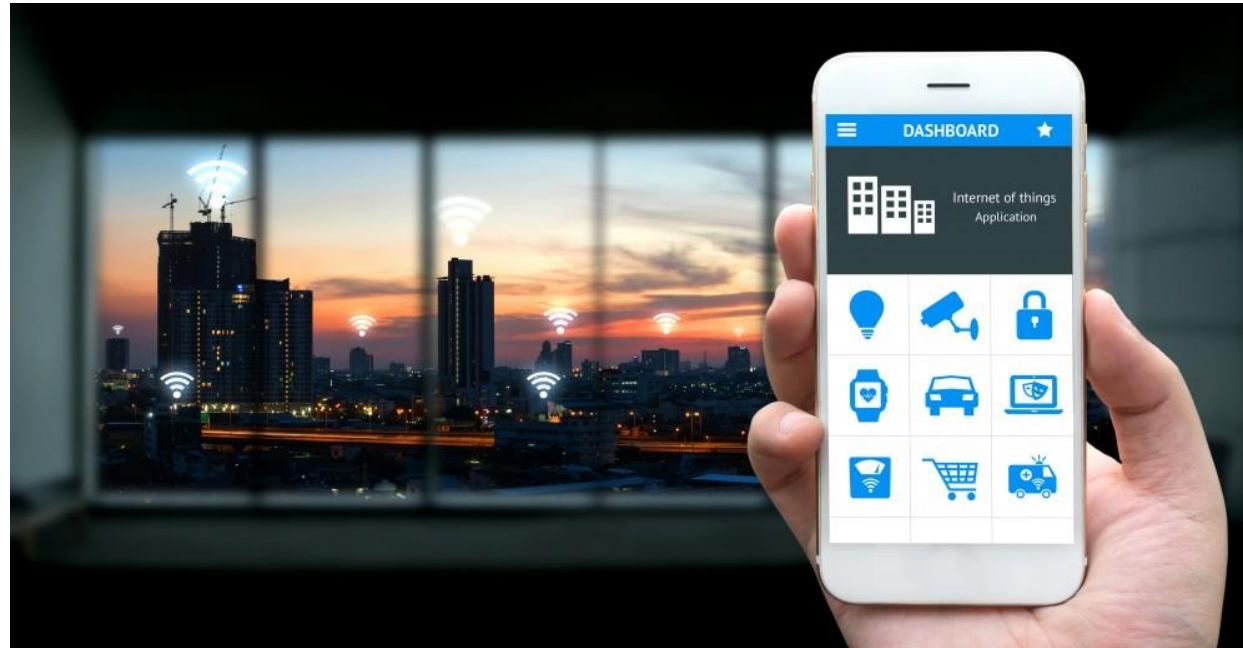
К примерам относятся решение задачи кредитного скоринга(оценки кредитоспособности), прогнозирование оттока клиентов, прогнозирование страховых рисков.

ИИ успешно применяется для прогнозирования спроса на товары и услуги, например рестораны могут с высокой точностью предсказать количество заказов на следующий день/дни и оптимизировать закупки продуктов, а курьерские службы и такси — выводить на линию оптимальное количество сотрудников.

К этому же направлению можно отнести рекомендательные системы — алгоритмы, прогнозирующие, какие товары/фильмы/продукты будут интересны клиенту, и маркетинговые инструменты, позволяющие предсказать, на какой баннер клиент более вероятно кликнет.



# Интернет вещей



Технология интернета вещей обещает в недалеком будущем войти в практически каждый дом: управлять техникой, оборудованием, а также через голосового помощника и мобильные приложения помогать пользователю планировать быт, досуг и работу

В такой экосистеме возникают самые разные задачи: некоторые чисто программистские — для которых известен алгоритм решения (к примеру, установить таймер или выключить свет), некоторые связаны с системами связи и электросетей (обеспечение доступа в интернет, вычислений), некоторые — более интеллектуальные, к примеру найти ответ на запрос пользователя, спланировать управление отоплением для поддержания заданной температуры или сформировать корзину для интернет-покупки недостающих в холодильнике продуктов.

Решение последних упомянутых интеллектуальных задач обычно строится через решение цепочки небольших подзадач: определить и посчитать товары на фотографии холодильника, спрогнозировать запрос пользователя на товары на ближайшую неделю, спланировать частоту покупки товаров, оптимизировать бюджет пользователя. Для каждой такой подзадачи разрабатывают отдельный алгоритм.

# Условия применения ИИ

Что необходимо для внедрения искусственного интеллекта в бизнес, какие условия для этого нужны.

## Данные

Машинное обучение построено на данных, без них не получится настроить алгоритм. Конечно, существует альтернатива — экспертные системы, в которых алгоритм предсказания реализуется на основе знаний специалиста (эксперта), однако на практике такие системы обычно показывают значительно более низкий уровень качества, чем алгоритмы машинного обучения. Если речь идет об автоматизации или прогнозировании процесса, который давно существует в компании (колл-центры, кредитный скоринг, прогнозирование спроса), то за это время, скорее всего, накопилось огромное количество данных, которые можно использовать для обучения. Известно, что необходимым условием в компании для

В таких случаях можно попытаться найти открытые данные в интернете или купить/заказать данные. Другая возможная проблема — безопасность: данные могут в компании быть, но выдать доступ разработчикам к ним нельзя. О решении подобных проблем стоит задуматься до начала работы над проектом. Также важен вопрос количества данных: чем их больше, тем точнее работает алгоритм, и чем сложнее задача, тем больше нужно данных.

## Ресурсы

Обучение моделей машинного обучения часто требует больших вычислительных мощностей. Настроить модель на относительно небольших данных можно на персональном компьютере, но для обработки данных о миллионах клиентов потребуются специальный сервер, причем даже с наличием сервера обработка может занимать дни или недели. Крупные компании часто покупают такие серверы (и нанимают специальную команду для их поддержки), небольшие компании могут арендовать облачные серверы, доступные через интернет.

## Специалисты

Для применения машинного обучения нужен опыт, поэтому настраивать алгоритм лучше доверить специалисту в этой области. Крупные компании обычно нанимают собственный штат, также существует множество вариантов для найма сторонней компании. В последнем случае заказчик снабжает исполнителя данными, описывает суть задачи и устанавливает метрики, а исполнитель занимается построением модели (и, возможно, внедрением).

С другой стороны, построить несложную модель на небольших данных может даже новичок, для этого достаточно пройти базовый онлайн-курс. В данной ситуации важно, чтобы данные были однородными, качественными и без ошибок. Также важно понимать, что специалисты в машинном обучении, как правило, не очень хорошо знают прикладную область, и для успеха проекта крайне желательно, чтобы заказчик понимал основные принципы и границы применимости машинного обучения — тогда двум сторонам будет проще найти общий язык и решить возникающие проблемы.

## Метрики

Заказчик должен четко понимать, каковы цели решения задачи, какие метрики необходимо оптимизировать (точность предсказания, удовлетворенность клиентов, доход) и какие значения метрики приемлемы с точки зрения бизнеса.

Бывает, что достижение хороших показателей на метриках машинного обучения не приводит к прорыву в бизнесе. Кроме того, достижение желаемого уровня метрик может быть слишком дорогим (требовать работы команды в течение нескольких месяцев).

Согласно некоторым данным, около половины проектов, связанных с искусственным интеллектом, терпят неудачу, среди причин выделяют нереалистичные ожидания, нехватку данных и проблемы с данными.

К сожалению, даже в случае выполнения указанных пунктов (имеется большое количество данных, ресурсов, и алгоритм настроен максимально качественно), проект может завершиться неудачей, не достигнув желаемых значений метрик качества.

В этом случае говорят об отсутствии зависимости в данных: когда по входу невозможно предсказать выход. Такие случаи возникают, например, на бирже (значения индексов зависят от множества факторов, и их крайне сложно предсказать). Однако даже в биржевой торговле сегодня активно применяются алгоритмы ИИ.