

Что такое искусственный интеллект



Разумный робот Дзенъятта. Скриншот из игры Overwatch. © Blizzard.

Тем, кто только начинает свой путь в изучении искусственного интеллекта (ИИ, ИскИн, *Artificial Intelligence*, AI), подчас бывает сложно разобраться с тем, что это вообще такое. Несмотря на то, что в окружающем инфополе этот термин встречается довольно часто, помощи в понимании это не добавляет, а иногда и просто вредит. Проблема в том, что практически везде он трактуется по-разному.

В этой статье мы проведем обзор существующих определений ИИ, попытаемся их систематизировать и разложить полученные знания по полочкам.

Хаос

Наше исследование начнем с поиска и анализа существующих определений ИИ, а первым источником, к которому мы обратимся, будет Википедия.

(1) Так в русскоязычной версии Вики [термин ИИ](#) определяется следующим образом:

Искусственный интеллект — свойство интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека.

(2) В англоязычной версии [этот же термин](#) определен несколько по другому:

Artificial intelligence is intelligence demonstrated by machines, as opposed to natural intelligence displayed by animals including humans.

(перевод на русский)

*Искусственный интеллект — это **интеллект, демонстрируемый машинами**, в отличие от естественного интеллекта, демонстрируемого животными, включая людей.*

По факту мы видим, что в одной статье термин ИИ определяется через свойство системы, а в другой через противопоставление естественного и машинного интеллектов. Причем оба вида интеллекта рассматриваются через призму поведения.

(3) В книге Романа Душкина «Искусственный интеллект» [1] термин ИИ определяется следующим образом:

*«... искусственный интеллект – это междисциплинарная **область исследований** и набор технологий, позволяющий создавать технические системы, решающие задачи, ранее доступные только человеку. ...»*

(4) [В Большой российской энциклопедии термин](#) ИИ определен так:

*Искусственный интеллект, **раздел информатики**, в котором разрабатываются методы и средства компьютерного решения интеллектуальных задач, традиционно решаемых человеком.*

В двух последних рассмотренных источниках мы видим, что авторы рассматривают ИИ не как техническую систему, а как науку, изучающую подобные системы.

(5) В книге «Искусственный интеллект: современный подход» [2] авторы тоже попытались сформулировать определение ИИ (правда, потом отказались от этой идеи и выбрали другой подход). Они, также как и мы сейчас, провели поиск определений ИИ, после чего свели в таблицу найденные результаты:

Таблица 1.1. Некоторые определения искусственного интеллекта, распределенные по четырем категориям

Системы, которые думают подобно людям	Системы, которые думают рационально
“Новое захватывающее направление работ по созданию компьютеров, способных думать, ...машин, обладающих разумом, в полном и буквальном смысле этого слова” [631]	“Изучение умственных способностей с помощью вычислительных моделей” [239]
“[Автоматизация] действий, которые мы ассоциируем с человеческим мышлением, т.е. таких действий, как принятие решений, решение задач, обучение...” [95]	“Изучение таких вычислений, которые позволяют чувствовать, рассуждать и действовать” [1603]
Системы, которые действуют подобно людям	Системы, которые действуют рационально
“Искусство создания машин, которые выполняют функции, требующие интеллектуальности при их выполнении людьми” [871]	“Вычислительный интеллект — это наука о проектировании интеллектуальных агентов” [1227]
“Наука о том, как научить компьютеры делать то, в чем люди в настоящее время их превосходят” [1285]	“Искусственный интеллект... — это наука, посвященная изучению интеллектуального поведения артефактов ¹ ” [1146]

Примечание. За расшифровкой ссылок из таблицы прошу обращаться в первоисточник.

В этих определениях ИИ рассматривался скорее с позиции области исследования, нежели с позиции технической системы.

Книги книгами, но давайте теперь посмотрим на что-то более официальное. В рамках Росстандарта работает технический комитет [\(ТК 164\) «Искусственный интеллект»](#), который занимается стандартизацией в интересующей нас области. Проанализируем плоды его трудов.

(6) [ГОСТ Р 59276-2020 «Национальный стандарт Российской Федерации. Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения»](#)

3.6 искусственный интеллект, ИИ: Способность технической системы имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных практически значимых задач обработки данных результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека.

(7) [ГОСТ Р 59277-2020 «Системы искусственного интеллекта. Классификация систем искусственного интеллекта»](#)

3.18 искусственный интеллект (artificial intelligence): Комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение, поиск решений без заранее заданного алгоритма и достижение инсайта) и получать при выполнении конкретных практически значимых задач обработки

данных результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Примечание — Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе, в котором используются методы машинного обучения). процессы и сервисы по обработке данных, анализу и синтезу решений.

ГОСТы смотрят на ИИ с позиции технической системы. Первый стандарт определяет ИИ как «способность технической системы...», что по сути характеризует поведение, а второй определяет ИИ как «комплекс технологических решений...», что в общем-то указывает на состав системы. На самом деле очень удивило, что в двух, выпущенных одним и тем же комитетом и утверждённых в один и тот же день стандартах используются разные определения. Ну да ладно, с отечественных стандартов переведем взор на международные.

(8) В проекте стандарта [ISO/IEC DIS 22989 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology](#) термин ИИ определен так:

3.1.2 artificial intelligence AI (engineered system) set of methods or automated entities that together build, optimize and apply a model (3.1.26) so that the system can, for a given set of predefined tasks (3.1.37), compute predictions (3.2.12), recommendations, or decisions

Note 1 to entry: AI systems are designed to operate with varying levels of automation (3.1.7).

Note 2 to entry: Predictions (3.2.12) can refer to various kinds of data analysis or production (including translating text, creating synthetic images or diagnosing a previous power failure). It does not imply anteriority.

3.1.3 artificial intelligence AI (discipline) study of theories, mechanisms, developments and applications related to artificial intelligence (engineered system) (3.1.2)

Официальный перевод этих определений на русский язык можно взять из [проекта ГОСТа](#) и звучит он следующим образом:

2.1.2 искусственный интеллект (ИИ) (artificial intelligence, AI): Способность приобретать, обрабатывать, создавать и применять знания (2.1.5), определенные в форме модели (2.1.6), для выполнения одной или нескольких поставленных задач (2.1.7). Примечание — Данный термин приведен в контексте системы.

2.1.3 искусственный интеллект (artificial intelligence, AI): Дисциплина о

создании и изучении ИИ (2.1.2). Примечание — Данный термин приведен в контексте инженерной дисциплины.

Интересно то, что анализируемый ISO разделяет точку зрения на ИИ как на техническую систему и точку зрения на ИИ как на научную дисциплину. Техническая система при этом определяется через демонстрируемые ей свойства.

Со стандартами все, перейдем на законодательный уровень. Практически в каждой стране есть понимание того, что ИИ — это ключевое направление ближайшего будущего. Все страны хотят его развивать, чтобы стать первыми и самыми крутыми по этому вопросу. Свой взгляд на развитие ИИ страны, как правило, оформляют в виде национальных стратегий развития ИИ. Хороший обзор подобных стратегий можно посмотреть на [Medium](#)'е. Мы же с вами глянем лишь несколько наиболее значимых стратегий и посмотрим, как в них определяется термин ИИ.

(9) В [Указе Президента РФ от 10.10.2019 N 490 «О развитии искусственного интеллекта в Российской Федерации» \(вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года»\)](#) термин ИИ определен практически идентично определению, данному в ГОСТ Р 59277-2020.

(10) В американском законе [H.R.6216 — National Artificial Intelligence Initiative Act of 2020](#) под ИИ понимают следующее:

The term “artificial intelligence” means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to—

(A) perceive real and virtual environments;

(B) abstract such perceptions into models through analysis in an automated manner; and

(C) use model inference to formulate options for information or action.

(перевод на русский)

Термин ‘искусственный интеллект’ означает техническую систему, которая для определенных человеком целей может делать прогнозы, выдавать рекомендации или принимать решения, влияющие на реальную или виртуальную среду. Системы искусственного интеллекта используют машинные и предоставленные человеком входные данные для:

(A) восприятия реальной и виртуальной среды;

(В) автоматического абстрагирования подобного восприятия в модели;

(С) использования полученных моделей для формирования информации или выполнения действий.

(11) Европейский союз сформулировал свое видение развития ИИ в документе [EUROPEAN COMMISSION. COMMUNICATION FROM THE COMMISSION. Artificial Intelligence for Europe {SWD\(2018\) 137 final}](#). В частности под ИИ европейцы понимают следующее:

Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals.

(перевод на русский)

Искусственный интеллект (ИИ) относится к системам, демонстрирующим разумное поведение, которое заключается в анализе окружающей среды и совершении действий – с некоторой степенью автономии – для достижения поставленных целей.

Для систематизации всего этого хаоса определений распишем учебную модель предполагаемого ИИ, к которой затем применим найденные определения и посмотрим, что из этого получится.

Модель

Как несложно было догадаться, в качестве модели будет выступать разумный робот-монах [Дзенъятта, путешествующий по миру в поисках духовного просветления](#).

В отличие от всяких «терминаторов» Дзенъятта — добрый и призван лечить союзников. С технической точки зрения в нем можно выделить следующие составные части:

1. **Сенсоры:** видеокамеры, микрофоны и другие датчики, с помощью которых робот получает информацию об окружающей среде и своем состоянии.
2. **Исполнительные механизмы:** руки, ноги, динамики и т.д — то есть компоненты робота, с помощью которых он воздействует на окружающую среду или изменяет свое состояние.

3. **Вспомогательные системы:** вычислительные устройства, на базе которых функционирует система управления, источник энергии, система коммутации, каркас, система смазки и т.д.

Схема информационных потоков, циркулирующих в роботе, выглядит следующим образом:



1. Данные о состоянии среды (внутреннее состояние робота будем далее считать частью среды) оцифровываются сенсорами и передаются в систему управления робота.
2. Система управления в целях решения стоящих перед роботом задач генерирует управляющие команды на основании сенсорных данных.
3. Управляющие команды поступают в исполнительные механизмы, в результате чего те выполняют определенные действия, осуществляя воздействие на среду.

Применяем определения к модели

Что в нашей модели будет является искусственным интеллектом?

По определениям (1), (2), (6), (11) ИИ — это не какая-либо часть робота, а **особенности его поведения**, проявляемые им в ходе функционирования.

Определения (3), (4), (5) рассматривают ИИ как **область исследования** и не применимы к технической системе, которой является робот.

Определения (7), (8) и (10) рассматривают ИИ как **«софт и железо», на котором он функционирует**. Соответственно в нашем случае к ИИ можно отнести всего робота целиком.

Классная статья получилась! Вроде в начале обещали разложить все по полочкам, а в итоге вышло непонятно что. К сожалению, так и есть, и с определениями ИИ творится действительно непонятно что, вернее понятно — полный бардак.

Подобный расклад происходит по множеству причин, и самая главная заключается в том, что ИИ до сих пор не изобрели. Тот, кто это сделает, тот и даст ему единое правильное толкование. Но почему его еще не изобрели? Во многом потому, что сам термин очень неудачный.

Что не так с терминологией ИИ

Основной проблемой с термином ИИ является использование в его трактовке человека (человеческого интеллекта, человеческого поведения, человеческих когнитивных способностей и т.д.). Конечно, понятно, почему так происходит — в глубине души люди мечтают разработать искусственную разумную жизнь по своему образу и подобию. А поскольку считается, что разумом обладает только человек, то и все определения делаются на основании сравнения ИИ с человеческим разумом. Однако, понятие «человеческий интеллект (разум)» в технических терминах не определено, и нет никакой уверенности, что его вообще можно определить.

Кроме того, говоря про разум или разумное поведение, часто имеют в виду поведение нормального здорового человека, мотивированного в решении поставленной задачи. Ок, а что, если поведение человека не вписывается в это правило? Например, человек болен шизофренией в стадии дебильности или всячески саботирует решение стоящей перед ним задачи. Получается, что у подобных людей нет интеллекта, но ведь это не так.

Еще одна проблема в том, что нет точной уверенности, что человеческий разум можно представить в виде алгоритма или программы. Ведь, если это так, то никакой пресловутой свободы воли нет, и человек по сути не решает ничего, все его действия предрешены и обусловлены его внутренним состоянием и состоянием окружающей среды. Но это уже религиозный спор, развивать его не будем.

Если бы люди пытались в лоб повторить полет бабочки, то авиации у человечества не было бы до сих пор. Вывод: для описания «разумных машин» следует использовать термины, не содержащие в себе отсылки к человеку.

Одним из шагов, направленных на то, чтобы в исследовании ИИ уйти от сравнения с человеком, послужило введение понятия **агент**. В проекте международного стандарта ISO/IEC DIS 22989 этому понятию дается следующая трактовка:

3.1.1 agent — automated entity that perceives its environment and takes actions to achieve its goals

Note 1 to entry: An AI agent is an agent that maximizes its chance of successfully achieving its goals by using AI techniques.

(русский перевод)

3.1.1 агент — автоматизированный объект, который воспринимает свое окружение и предпринимает действия для достижения своих целей. Примечание 1 к записи: Агент ИИ — это агент, который максимизирует свои шансы на успешное достижение своих целей с помощью методов ИИ.

В книге «Искусственный интеллект: современный подход» [2] термин агент определяют так:

«Агентом считается все, что действует. Но предполагается, что компьютерные агенты обладают некоторыми другими атрибутами, которые отличают их от обычных „программ“, такими как способность функционировать под автономным управлением, воспринимать свою среду, существовать в течение продолжительного периода времени, адаптироваться к изменениям и обладать способностью взять на себя достижение целей, поставленных другими.»

Как отмечалось ранее, авторы этой книги разочаровались в термине ИИ как объекте исследования и выбрали вместо него термин «рациональный агент».

Стоит отдельно обратить внимание на то, что агентом можно считать не только «разумного» робота, возможны и другие варианты, например:

1. Игровой бот — программа, управляющая действиями персонажей в компьютерных играх. Сенсоры бота — это программный код, получающий состояние среды от игры через ее API. Исполнительные механизмы — команды API игры, приводящие к изменению поведения управляемого персонажа. Если реализация ботов изначально не заложена создателями игры, то сенсорами могут выступать программные модули, перехватывающие вывод на экран, а исполнительными механизмами — программный код, эмулирующий нажатия кнопок на клавиатуре и движения курсора мыши.

2. Интеллектуальная система управления дорожным движением.
Сенсоры — видеокамеры на дорогах, исполнительные механизмы — светофоры.

3. и др.

Идея. Интересным объектом исследования может послужить компьютерный вирус, построенный по принципам рационального агента.

Сила интеллекта



(с) Яндекс.Картинки

До этого момента мы рассматривали термин ИИ только в контексте «разумного» робота, вернее агента, но в инфополе данный термин используется шире. Например, к технологиям ИИ относят разработки в

области распознавания речи, синтеза изображений по описанию и др.

Тут следует дать пояснение. Дело в том, что все разработки в области ИИ условно разделяют на несколько категорий: [слабый ИИ](#) и [сильный ИИ](#). Простыми словами можно сказать, что слабый ИИ проявляет ограниченные возможности и позволяет решать узкоспециализированные задачи, которые раньше мог решать только человек. Сильный ИИ, в отличие от слабого, обладает всеми возможностями человеческого разума, а возможно и лучше (причем это «лучше» в итоге может привести к [технологической сингулярности](#)). Различные варианты слабого ИИ, такие как голосовые помощники, автопилоты и др., прочно вошли в нашу жизнь, в то время как сильный ИИ — это пока не достижимая мечта. Ну и рассматриваемый нами разумный робот — это, конечно, сильный ИИ. Не исключено, что сильный ИИ будет сделан в результате удачного объединения нескольких технологий слабого ИИ. Далее под ИИ будем понимать только сильный ИИ.

Свои пять копеек

Какая же статья про ИИ обойдется без попытки автора определить этот термин своими словами? Наверняка такие статьи есть, но эта к ним точно не относится. Не буду растекаться мыслью по древу и максимально кратко изложу свое видение этого вопроса.

1. Человеческий интеллект — это то, что управляет поведением человека, соответственно ИИ — это то, что управляет поведением агента, таким образом ИИ по сути является системой автоматического управления (САУ).
2. Отличительными особенностями ИИ от других видов САУ будет реализация ей следующих свойств:
 1. ...
 2. ИИ — это система, совершенствующая сама себя (или по-другому: ИИ — это программа, пишущая сама себя).

Изначально здесь я описал довольно большое количество свойств, среди которых были автономность, адаптация, управление знаниями и т.д. Потом подумал и удалил их. При желании о них можно узнать из других источников. Тут же я хочу отметить лишь одну ключевую, на мой взгляд, характеристику системы, придающую ей интеллектуальность — это самосовершенствование.

Данное качество направлено на повышение эффективности решения стоящих перед агентом задач. Оно может проявляться в уменьшении ресурсов, требуемых для решения задачи, повышения скорости

достижения целей и т.д. Совершенствование ИИ заключается в виде совершенствования моделей данных, с помощью которых он описывает среду функционирования, и совершенствование (или создание новых) алгоритмов, используемых для решения задач.

Если сравнивать предлагаемое качество с самообучением, речь о котором шла в определении ИИ в ранее рассмотренных ГОСТах, то самосовершенствование несколько шире и на, мой взгляд, лучше отражает интеллектуальность.

Резюме

1. С определениями ИИ беда — единого трактования нет, практически все под ИИ подразумевают что-то свое.
2. Большую часть найденных в процессе исследования трактовок термина ИИ можно разделить на два больших класса: первый — рассматривающий ИИ в контексте технической системы; второй — рассматривающий ИИ в контексте научной дисциплины.
3. Во многих трактовках термина ИИ используется сравнение с человеком (человеческим интеллектом, человеческими когнитивными способностями и т.д.). Данный факт вносит дополнительную неопределенность, поскольку все, что связано с человеком, крайне сложно или даже невозможно определить в технических терминах.
4. В некоторых источниках делаются попытки трактовать термин ИИ без отсылки к человеку (например, в ISO/IEC DIS 22989). В других источниках при попытке изучать «разумное поведение машин» делается попытка уйти от термина ИИ и перейти к термину рациональный агент [2].
5. Автор статьи предлагает считать, что ИИ это одна из разновидностей систем автоматического управления, а ее главной (но не единственной) отличительной способностью является возможность самосовершенствования.

Литература

1. Душкин Р. В. Искусственный интеллект. — М.: ДМК Пресс, 2019. — 280 с., ISBN 978-5-97060-787-9.
2. Рассел, Стюарт, Норвинг, Питер. Искусственный интеллект: современный подход, 2-е изд.: Пер. с англ. — М.: Издательский дом «Вильямс», 2006. — 1408 с.: ил. — Парал. тит. англ. ISBN 5-8459-0887-6 (рус.)