

# Содержание

|   |           |
|---|-----------|
| Об авторе .....   | 11        |
| О соавторе .....  | 12        |
| О техническом рецензенте .....  | 13        |
| Признательности .....   | 14        |
| Введение .....  | 15        |
| Примечания переводчика .....  | 18        |
| <b>Глава 1. Начало работы с основами ИИ .....</b>                               | <b>19</b> |
| Правда и вымысел .....  | 20        |
| История и эволюция .....  | 21        |
| Нынешнее состояние дел .....  | 23        |
| Превращение ИИ в товар .....  | 24        |
| Microsoft и ИИ .....  | 25        |
| Фундаментальные понятия .....   | 26        |
| Машинное обучение .....   | 28        |
| Язык .....  | 31        |
| Речь .....  | 33        |
| Компьютерное зрение .....   | 34        |
| Когнитивные службы Microsoft .....  | 37        |
| Зрение .....  | 38        |
| Речь .....  | 39        |
| Язык .....  | 39        |
| Знание .....  | 40        |
| Поиск .....   | 41        |
| Резюме .....  | 41        |
| <b>Глава 2. Создание приложения с поддержкой ИИ в среде Visual Studio .....</b> | <b>43</b> |
| Необходимые условия для использования Когнитивных служб .....                   | 44        |
| Подготовка среды разработки .....   | 44        |
| Получение ключа подписки Azure для Когнитивных служб .....                      | 45        |
| Шаг 1. Зарегистрировать учетную запись Azure .....                              | 45        |
| Шаг 2. Создать новую учетную запись Когнитивных служб .....                     | 48        |
| Шаг 3. Получить ключ (ключи) подписки .....                                     | 50        |
| Тестирование API .....  | 51        |
| Что вы хотите сделать .....   | 52        |
| Как это сделать .....   | 52        |
| Создание первого ИИ-ориентированного приложения .....                           | 55        |
| Исходный код .....  | 56        |
| Пошаговый анализ .....  | 57        |
| Результат .....   | 61        |
| Более интересный вариант приложения .....                                       | 61        |

|                                       |    |
|---------------------------------------|----|
| Извлечение текста из изображений..... | 61 |
| Исходный код .....                    | 62 |
| Пошаговый анализ .....                | 64 |
| Результат .....                       | 64 |
| Резюме.....                           | 65 |

|  |           |
|--|-----------|
| <b>Глава 3. Конструирование разговорного интерфейса с применением технологий Microsoft .....</b> | <b>66</b> |
| Что такое разговорный интерфейс пользователя? .....  | 68        |
| Краткая история .....  | 69        |
| В самом начале: интерфейс командной строки.....  | 69        |
| И затем появился графический интерфейс пользователя .....  | 71        |
| Очередная эволюция: разговорный интерфейс пользователя .....                                     | 72        |
| Роль ИИ в разговорном интерфейсе пользователя.....   | 75        |
| Ловушки разговорного интерфейса пользователя .....   | 76        |
| Будущее может быть за гибридным интерфейсом пользователя<br>(разговорный + графический) .....    | 78        |
| Принципы разработки .....  | 79        |
| Инфраструктура Microsoft Bot Framework.....  | 80        |
| Создание приложения с разговорным интерфейсом в среде Bot Framework .....                        | 81        |
| Шаг 0. Подготовка среды разработки.....  | 82        |
| Шаг 1. Создание нового проекта приложения Bot Framework.....                                     | 82        |
| Шаг 2. Первое сообщение и сообщение по умолчанию .....   | 83        |
| Шаг 3. Выполнение, тестирование и отладка вашего бота .....                                      | 85        |
| Шаг 4. Планирование визита к врачу .....   | 88        |
| Шаг 5. Обработка системных сообщений .....   | 91        |
| Следующие шаги .....   | 92        |
| Резюме.....  | 92        |

|  |           |
|--|-----------|
| <b>Глава 4. Применение технологии понимания естественного языка.....</b> | <b>93</b> |
| Что такое понимание ЕЯ? .....  | 94        |
| История развития методов понимания естественного языка .....             | 97        |
| Почему машины с трудом понимают естественный язык .....                  | 100       |
| Сложности естественного языка .....                                      | 100       |
| Статистические модели в качестве решения недостаточны.....               | 102       |
| Многообещающее будущее .....   | 104       |
| Облачная служба LUIS .....   | 105       |
| Архитектура программного приложения на основе LUIS .....                 | 106       |
| За кадром .....  | 109       |
| Масштабная тренировка – основа всему .....                               | 110       |
| Получение подписки Azure для облачной службы LUIS .....                  | 112       |
| Демонстрационный пример: приложение DefinitionApp .....                  | 114       |
| Примечания .....   | 116       |
| Резюме.....  | 117       |

|  |            |
|--|------------|
| <b>Глава 5. Исследование когнитивной модели языка.....</b> | <b>118</b> |
| API Bing проверки орфографии .....                         | 118        |
| Что это такое? .....                                       | 121        |
| Как его использовать .....                                 | 122        |

|   |            |
|---|------------|
| Интеграция с облачной службой LUIS .....                            | 125        |
| API анализа текста .....  | 127        |
| Идентификация языка .....   | 127        |
| Выделение ключевых фраз .....                                       | 131        |
| Анализ мнений .....   | 133        |
| Идентификация темы .....  | 135        |
| Идеи использования .....  | 138        |
| API веб-модели языка (WebLM) .....                                  | 139        |
| Разбиение на слова .....  | 141        |
| Совместная вероятность .....  | 142        |
| Условная вероятность .....  | 143        |
| Предсказание следующего слова .....                                 | 145        |
| API лингвистического анализа .....                                  | 146        |
| Разбиение на предложения и лексемизация .....                       | 147        |
| Частеречная разметка .....  | 150        |
| Синтаксический анализ на основе грамматики составляющих .....       | 152        |
| Резюме .....  | 154        |
| <b>Глава 6. Потребление и применение облачной службы LUIS .....</b> | <b>155</b> |
| Планирование своего приложения .....                                | 156        |
| Что бот должен уметь делать? .....                                  | 156        |
| Какая информация нужна боту от пользователя? .....                  | 157        |
| Что должно быть сделано в облачной службе LUIS? .....               | 157        |
| Что должно быть сделано в Bot Framework? .....                      | 158        |
| Создание приложения с поддержкой LUIS .....                         | 159        |
| Добавление намерений .....  | 160        |
| Добавление/разметка высказываний .....                              | 160        |
| Публикация своего приложения .....                                  | 162        |
| Добавление сущностей .....  | 164        |
| Простые сущности .....  | 166        |
| Составные сущности .....  | 168        |
| Иерархические сущности .....  | 171        |
| Предварительно подготовленные сущности .....                        | 173        |
| Добавление списка фраз .....  | 174        |
| Рекомендуемые последующие шаги .....                                | 175        |
| Активное обучение посредством рекомендованных высказываний .....    | 175        |
| Использование API LUIS для автоматизации .....                      | 176        |
| Интеграция LUIS со средой Bot Framework .....                       | 176        |
| Создание проекта в Visual Studio .....                              | 177        |
| Обработка намерения без сущностей .....                             | 178        |
| Подготовка своего бота для использования HealthCheckupDialog .....  | 178        |
| Тестирование бота в эмуляторе .....                                 | 179        |
| Обработка намерения с полным комплектом сущностей .....             | 179        |
| Обработка намерения с составными сущностями .....                   | 181        |
| Обработка намерения None .....                                      | 183        |
| Добавление бота в Skype .....                                       | 184        |
| Публикация своего бота .....  | 184        |
| Регистрация своего бота .....                                       | 184        |
| Резюме .....  | 186        |

|   |     |
|---|-----|
| <b>Глава 7. Взаимодействие с API обработки речи</b> .....               | 187 |
| Способы взаимодействия с речью.....                                     | 188 |
| API когнитивного поиска.....  | 189 |
| Распознавание речи.....   | 189 |
| Начало работы.....  | 190 |
| Сначала получение JWT-токена.....                                       | 190 |
| API персонализированной обработки речи.....                             | 192 |
| Синтез речи.....  | 193 |
| Механизм распознавания речи.....  | 196 |
| Служба персонализированной обработки речи.....                          | 197 |
| Персонализированная акустическая модель.....                            | 197 |
| Персонализированная модель языка.....                                   | 207 |
| Данные произношения.....  | 209 |
| Персонализированная конечная точка преобразования речи в текст.....     | 210 |
| Распознавание говорящего.....   | 213 |
| Верификация говорящего против идентификации говорящего.....             | 213 |
| Регистрация для верификации говорящего.....                             | 214 |
| Верификация говорящего.....   | 217 |
| Регистрация для идентификации говорящего.....                           | 217 |
| Идентификация говорящего.....   | 219 |
| Состояние операции.....   | 219 |
| Резюме.....   | 220 |
| <br>  |     |
| <b>Глава 8. Применение поисковых предложений</b> .....                  | 221 |
| Поиск повсюду.....  | 222 |
| Повсеместность, предсказательность, проактивность (три «П» поиска)..... | 222 |
| История Bing.....   | 224 |
| В чем уникальность Bing?.....   | 226 |
| API поиска.....   | 226 |
| API Bing автозаполнения.....  | 226 |
| Как использовать API Bing автозаполнения.....                           | 227 |
| API Bing поиска изображений.....  | 231 |
| Как использовать API Bing поиска изображений.....                       | 232 |
| API Bing поиска новостей.....   | 237 |
| API Bing поиска видео.....  | 240 |
| Как использовать API Bing поиска видео.....                             | 240 |
| API Bing поиска в сети.....   | 244 |
| Как использовать API Bing поиска в сети.....                            | 244 |
| Резюме.....   | 248 |
| <br>  |     |
| <b>Глава 9. Работа с рекомендациями</b> .....                           | 249 |
| Понимание основ.....  | 250 |
| Рекомендация товаров, часто приобретаемых вместе.....                   | 251 |
| Рекомендация «от предмета к предмету».....                              | 252 |
| Рекомендация на основе предыстории.....                                 | 253 |
| Как эти рекомендации работают?.....                                     | 253 |
| Модели и типы рекомендаций.....   | 258 |
| Сборка обычной модели рекомендаций.....                                 | 259 |
| Сборка модели рекомендации на основе FBT.....                           | 262 |

|  |            |
|--|------------|
| Сборка модели рекомендаций на основе ранжирования..... | 265        |
| Сборка модели рекомендации на основе SAR .....         | 267        |
| Установление правил в сборке .....                     | 270        |
| Офлайнное оценивание .....                             | 270        |
| Веб-интерфейс API рекомендаций.....                    | 271        |
| Резюме.....  | 275        |
| <b>Глава 10. ИИ в ближайшем будущем .....</b>          | <b>276</b> |
| Почему ИИ имеет такую популярность?.....               | 277        |
| Рост вычислительной мощности .....                     | 277        |
| Открытия в алгоритмах ИИ .....                         | 278        |
| Данные – это новая валюта.....                         | 279        |
| Возникновение облачных вычислений.....                 | 280        |
| Службы против решений?.....                            | 281        |
| Когнитивные категории.....                             | 281        |
| Проблемы и будущее понимания ЕЯ.....                   | 282        |
| Проблемы и будущее обработки речи.....                 | 283        |
| Проблемы и будущее поиска .....                        | 284        |
| Проблемы и будущее рекомендаций .....                  | 284        |
| Прежде всего ИИ.....                                   | 285        |
| Интеллектуальная периферия.....                        | 285        |
| Устранение задач, а не рабочих мест .....              | 286        |
| Так куда же мы катимся? .....                          | 289        |
| <b>Тематический указатель .....</b>                    | <b>291</b> |

# Об авторе



Нишит Патак носит звание наиболее ценного специалиста компании Microsoft (MVP), является системным архитектором, лектором, а также идеологом, новатором и стратегом ИИ. Он является плодовитым писателем и автором и написал много книг, статей, обзоров и рубрик для нескольких электронных и печатных изданий. Более чем 20-летний опыт работы в области ИТ позволил Нишиту накопить экспертные знания, которые лежат в области инноваций, исследований, проектирования архитектуры, проектирования

и разработки приложений для компаний из списка 100 крупнейших промышленных компаний США с использованием инструментов и технологий следующего поколения. Являясь ранним приверженцем технологии Microsoft, он идет в ногу в проблемах, связанных с сертификацией программных продуктов, и преуспел в получении нескольких своих сертификатов на стадии бета-тестирования.

Нишит является золотым членом консультативного совета различных национальных и международных сообществ и организаций в области информатики. Он дважды был награжден элитным званием наиболее ценного специалиста компании Microsoft (MVP) за свою образцовую работу и опыт в технологиях Microsoft. Он является членом различных консультативных групп корпорации Microsoft. В настоящее время Нишит работает вице-президентом и руководителем НИОКР в Accenture Technology Labs. Он сосредоточен на ключевых областях исследований, в частности искусственном интеллекте, машинном обучении, когнитивных вычислениях, программах-роботах и облачных вычислениях на основе технологии блокчейна, а также помогает компаниям проектировать архитектуру решений на основе этих технологий. Нишит родился, вырос и получил образование в городе Котдвара в Уттаракханде, Индия. Когда позволяет время, он проводит его с семьей и друзьями и забавляет их своими познаниями в области хиромантии и астрологии. Вы можете с ним связаться по адресу [nispathak@gmail.com](mailto:nispathak@gmail.com).

# О соавторе



Анураг Бхандари является исследователем, программистом и проповедником свободного программного обеспечения. Его любимые области исследований включают обработку ЕЯ, интернет вещей и машинное обучение. Он специализируется на разработке веб- и мобильных приложений и решений. Имеет большой опыт работы с компаниями из списка 500 крупнейших промышленных компаний, стартапами и НКО в качестве исследователя и разработчика программного обеспечения. Анураг родом из Джаландхара, Пенджаб, где

он также получил диплом в области компьютерных наук в Национальном технологическом институте. Со своих студенческих дней он был связан с несколькими проектами с открытым исходным кодом, такими как Granular Linux и OpenMandriva, и в ряде из них возглавлял эти проекты. Он является потрясающим полиглотом в языках программирования (C#, Java, JavaScript, PHP, Python) и естественных языках (английский, хинди, панджаби, французский). Будучи энтузиастом в области технологий, Анураг продолжает возиться с трендовыми технологиями и испытывать новые программные инфраструктуры и платформы. В свободное время он читает книги, занимается спортом, мечтает о гаджетах, смотрит телепередачи, играет в игры и собирает марки. Вы можете найти его онлайн на <http://anuragbhandari.com> или оставить ему записку по адресу [anurag.bhd@gmail.com](mailto:anurag.bhd@gmail.com).

# О техническом рецензенте

Фабио Клаудио Ферракьятти является старшим консультантом и старшим аналитиком/разработчиком с использованием технологий Microsoft. Он работает в Blu-Arancio С. с. А ([www.bluarancio.com](http://www.bluarancio.com)) старшим аналитиком/разработчиком и специалистом Microsoft Dynamics CRM, является сертифицированным разработчиком решений Microsoft для .NET, дипломированным специалистом по программным средствам Microsoft, а также плодовитым автором и техническим редактором. За последние десять лет он написал ряд статей для итальянских и международных журналов и являлся соавтором более десяти книг на различные компьютерные темы.



# Признательности

Эта книга стала результатом коллективной работы нескольких замечательных людей. Она не была бы завершена без моего партнера Анурага Бхандари, который проделал фантастическую работу, помогая завершать главы, писать исходный код и проводить исследования. В разное время мы проговаривали темы, обсуждали технологии и формировали книгу в правильном направлении. Анураг, ты – «единственный человек», который помог мне поднять эту книгу далеко за пределы моих ожиданий.

Спасибо всем сотрудникам издательства Apress, которые приложили искренние усилия, чтобы опубликовать данную книгу. Гвенан заслуживает особой благодарности, потому что, прежде чем приступить к этому проекту, я долго общивался с ней письмами. Спасибо Нэнси и Лоре за отличную работу по управлению проектами и постоянное побуждение меня делать все возможное. Я также хотел бы поблагодарить Генри Ли за его технический обзор. Я бы не постеснялся сказать, что все вы – очень талантливые люди. Каждый из вас очень помог этой книге, и я с нетерпением жду возможности поработать с вами над следующей.

И последнее, но не менее важное: благодарю свою семью, в особенности жену Сурабхи и своего отца Панкаджа Патака, за то, что они были добры ко мне и всегда были готовы прийти на помощь, и за то, что мои мечты сбылись. Все, что я делаю в своей жизни, было бы невозможно без вас.

# Введение

Эта книга познакомит вас с миром искусственного интеллекта. В большинстве случаев разработчики считают задачу внедрения ИИ трудно выполнимой, так как она сопряжена с написанием сложных алгоритмов и сотен строк программного кода. Задача настоящей книги – показать, как с помощью созданного вами когнитивного приложения, состоящего из нескольких строчек программного кода, можно найти более простое решение. Это стало возможным благодаря появлению разнообразных прикладных программных интерфейсов Когнитивных веб-служб (Cognitive Services API). Данная книга посвящена нескольким полезным технологиям, благодаря которым ваше приложение сможет разумно применять API Когнитивных веб-служб. Компания Microsoft предоставляет разработчикам оптимальный опыт взаимодействия и реализует его с помощью своих программных интерфейсов Cognitive Services API.

Эта книга дает четкое представление о принципах ИИ. Темы поддержания естественно-языкового диалога, обработки естественного языка и выполнения интеллектуального поиска настолько многогранны, что для каждого из этих направлений потребовалась бы отдельная книга. В настоящей книге предпринимается попытка дать четкое объяснение по каждому из этих понятий, сначала давая ясные ответы на вопросы «Почему» и «Что» и только после этого вдаваясь в подробное объяснение того, «Как», т. е. каким образом устроен тот или иной API. В книге также приводятся подробные примеры, которые призваны облегчить применение новых концепций на практике. Книга «Искусственный интеллект для .NET: речь, язык и поиск» покажет, каким образом уже сегодня можно приступить к внедрению этих удивительных возможностей в свои собственные приложения.

Данная книга начинается с того, что знакомит вас с описанием и технологией искусственного интеллекта, проведя небольшой экскурс в его историю, терминологию и методологию. Затем вы познакомитесь с программными интерфейсами Когнитивных служб и соответствующими инструментами, после чего перейдете к пошаговому созданию вашего первого интеллектуального когнитивного приложения, используя для этого среду разработки Visual Studio. Далее познакомитесь с понятиями, связанными с разговорным интерфейсом пользователя (РИП, conversational user interface, CUI), и приступите к созданию своего первого бота на основе программной инфраструктуры Microsoft Bot Framework. Данная книга позволяет понять процесс планирования вашего приложения с использованием инфраструктуры Bot Framework, а также познакомит с лучшими практическими методами ее применения.

Настоящая книга также позволяет глубже познакомиться с проблемами понимания естественного языка (natural language understanding, NLU) и его обработки (natural language processing, NLP). Эти технологии позволяют компьютерным программам понимать людей точно так же, как люди понимают друг друга. Книга подробно останавливается на интеллектуальной облачной службе понимания естественного языка Microsoft LUIS (Language Understanding Intelligent Service) и принципах ее работы, а также на приемах проектирования, употребления

и применения службы LUIS перед созданием проекта с участием LUIS с нуля. Книга также содержит подробные инструкции по тестированию, тренировке и публикации приложения с участием LUIS перед его развертыванием и использованием в среде Bot Framework.

Речь – это самая естественная форма взаимодействия. Данная книга предоставляет пошаговое объяснение речевого программного интерфейса Speech API и приемов применения API для распознавания и синтеза речи. Далее вы подробно познакомитесь с примерами применения службы персонализированной обработки речи, ранее известной под названием CRIS и предоставляющей пошаговый план создания вашей первой модели языка, аудиомодели, ее развертывания, а также применения службы персонализированной обработки речи. В книге также содержится подробная информация о том, как узнать, кто сейчас говорит.

В 8-й главе книги вы найдете историю и подробное описание популярной поисковой системы Bing и интерфейса API и способы применения возможностей этого поисковика в своих приложениях. В книге также подробно рассказывается об основополагающих понятиях и типах рекомендаций и дается пошаговое описание способов получения этих рекомендаций. Книга заканчивается экскурсом в будущее искусственного интеллекта. Другими словами, эта книга может рассматриваться как путеводитель, руководствуясь которым, вы сможете осуществить ваши последующие шаги при создании своих приложений.

В этой книге вы:

- изучите основы технологии искусственного интеллекта на практических примерах и сценариях;
- начнете создавать приложение на основе ИИ в среде разработки Visual Studio;
- создадите текстоориентированный разговорный интерфейс для прямого взаимодействия с пользователем;
- примените API обработки речи Когнитивных служб для распознавания и интерпретации устной речи;
- ознакомитесь с различными моделями языка, включая обработку естественного языка, и примените их в приложении Visual Studio;
- повторно будете использовать поисковые возможности Bing, чтобы лучше понять, что именно пользователь имеет в виду;
- поработаете с предлагаемыми технологиями и внедрите их в свои приложения.

## Для кого эта книга предназначена

Термин «искусственный интеллект» представляет собой профессиональный жаргонизм современной информационной индустрии. Об ИИ говорят все. Из-за связанной с этим неразберихи разработчики программного обеспечения зачастую пребывают в недоумении по поводу того, где и как начинать работу с ИИ. И появление когнитивных программных интерфейсов Microsoft Cognitive API предлагает разработчикам широкий спектр новых функциональных возможностей. Эта книга предназначена для читателей начального и промежуточного уровней, интересующихся искусственным интеллектом. Разработчики и архитекторы программного обеспечения с предыдущим опытом работы в среде .NET или без та-

кового, которые хотят применить новые когнитивные API в своих приложениях, извлекут из предлагаемого в данной книге фактического материала и примеров программного кода немалую пользу. Эта книга также послужит отличным руководством для разработчиков приложений и архитекторов программного обеспечения, которые не знакомы с ИИ и/или ключевыми принципами применения некоторых когнитивных API.

## НЕОБХОДИМЫЕ КОМПОНЕНТЫ

Чтобы эта книга была наиболее полезна, вам потребуется программная платформа .NET Framework от компании Microsoft и подключение к Интернету. В качестве среды разработки для экспериментирования с примерами программного кода, которые можно найти в разделе Исходный код сайта Apress ([www.apress.com](http://www.apress.com)), рекомендуется использовать Microsoft Visual Studio 2017.

## ПОЛУЧЕНИЕ ОБНОВЛЕНИЙ ДЛЯ ЭТОЙ КНИГИ

По ходу чтения вы можете натолкнуться на случайную грамматическую или программную ошибку, хотя я думаю, что этого не произойдет. Если же такое случится, то примите мои искренние извинения. Каждый человек имеет право на ошибку, поэтому одна или две неточности могут присутствовать независимо от предпринятых мною усилий. Текущий список ошибок можно получить на веб-сайте Apress (который тоже находится на домашней странице этой книги). Здесь же можно узнать, как уведомить меня о любых ошибках, которые вы можете найти.

## СВЯЗЬ С АВТОРОМ

Если у вас есть какие-либо вопросы относительно исходного кода этой книги, или вы нуждаетесь в разъяснении того или иного примера, или же просто хотите предложить свои мысли относительно ИИ либо хотите связаться со мной по поводу других вопросов, то прошу, не стесняйтесь и напишите мне на [nispathak@gmail.com](mailto:nispathak@gmail.com). Я сделаю все возможное, чтобы своевременно с вами связаться.

Благодарю, что приобрели эту книгу. Надеюсь, что вы получите удовольствие от ее прочтения и примените свои вновь обретенные знания с пользой для всех.

## УСЛОВНЫЕ ОБОЗНАЧЕНИЯ



Такая пиктограмма обозначает совет.



Такая пиктограмма обозначает примечание.

# Примечания переводчика

Книга «Искусственный интеллект для .NET: речь, язык и поиск» является превосходным руководством по ИИ и технологиям Microsoft в этой области.

В переводе программный код адаптирован на основе Visual Studio 2017 Community и только в главах 2 и 3. В главах 4 и 5 используется интеллектуальная система понимания речи LUIS, в которой на сегодняшний день русский язык пока не используется, но его внедрение остается в планах. В последующих главах автор ссылается на программные шаблоны, созданные во второй и третьей главах.

По всей книге заменены ссылки на API с региона западных США на регион Западной Европы, который позволяет работать с API Когнитивных служб из России. К оригинальной версии исходного кода добавлена его адаптированная версия; добавлен шаблон запроса API Когнитивных служб на языке Python для блокнота Jupyter и в стандартном файле .py.

В целом предлагаемый материал имеет отличную практическую направленность и будет интересен русскоязычному читателю, лишь за небольшим исключением применения облачной службы LUIS, которая пока русский язык не поддерживает. Но наибольший интерес книга вызывает с точки зрения современных механизмов работы технологии ИИ вообще и в частности в Microsoft, где стремительно развивается комплект API Когнитивных служб (Cognitive Services API). В этом плане книга действительно является замечательным руководством. А примеры по созданию программы-бота в среде Bot Framework являются превосходным практическим дополнением.

# Глава 1

## Начало работы с основами ИИ

Представьте, что вы создаете программный продукт, который настолько умен, что понимает не только человеческие языки, но и жаргонные выражения и тончайшие диалекты этих языков, да так, что ваше программное обеспечение будет знать, что «Привет, компьютер! Как дела?» и «Как делишки, чувак?» означает одно и то же.

И пока вы этим занимаетесь, почему бы не добавить в ваше программное обеспечение возможность слушать говорящего человека и реагировать соответствующим образом?

*Пользователь:* «Какое у меня расписание на сегодня?»

*Софт:* У вас сегодня довольно насыщенный день, одна за другой встречи с 10 утра до 13:30 пополудни и потом снова с 15:00 до 19:00 вечера».

Но почему бы не добавить в данный продукт возможность вести разговоры, подобные человеческим диалогам? Тот факт, что софт может отвечать человеку, еще не обозначает, что сам программный продукт обладает разумом. Но почему бы не добавить в данный продукт возможность вести диалоги, подобные человеческим? Например.

*Пользователь:* «Компьютер, я пропустил матч? Какой счет?»

*Софт:* «Уже 31-я минута футбольного матча между “Барселоной” и “Реалом”. Ваша любимая команда “Барселона” еще не забила. Счет 0:1».

*Пользователь:* «Черт возьми! Кто забил у Реала?»

*Софт:* «Первый гол на 10-й минуте забил Криштиану Роналду».

*Пользователь:* «Мне это не нравится. Сколько он забил в этом сезоне?»

*Софт:* «На данный момент Роналду забил 42 гола за свой клуб и 13 голов за свою страну».

*Пользователь:* «Впечатляет. Надеюсь, бедняга Месси скоро догонит».

*Пользователь:* «Компьютер, спасибо за новости».

*Софт:* «Не за что».

*Софт:* «Не забудьте узнать счет через полчаса. Если опираться на статистику владения мячом и ударов по воротам, то есть 73%-й шанс, что Барселона в ближайшие 20 минут забьет гол».

Разве такие возможности не сделают ваше программное обеспечение *умным и интеллектуальным*? Являясь разработчиком в среде .NET, вы вскоре узнаете, каким образом *вы* могли бы сделать ваше программное обеспечение таким же умным, что и Microsoft Cortana, Apple Siri или Google Assistant.

Microsoft Cortana – это голосовая помощница Кортана от компании Microsoft.

Apple Siri – персональный голосовой помощник с искусственным интеллектом, присутствующий во всех современных аппаратах Apple.

Google Assistant – умный виртуальный ассистент от компании Google.

После прочтения этой главы вы:

- узнаете правду и вымысел об ИИ;
- познакомитесь с историей и эволюцией ИИ;
- познакомитесь с разработками Microsoft в области ИИ;
- познакомитесь с основными понятиями ИИ:
  - когнитивным машинным обучением, глубоким обучением, обработкой естественного языка (ЕЯ), пониманием ЕЯ и др.;
  - показательными диаграммами и ссылками (где это возможно);
- познакомитесь с Когнитивными службами Microsoft:
  - поговорим о всех пяти когнитивных группах;
- узнаете, как применять ИИ в своем программном обеспечении;
- узнаете о перспективах развития ИИ в обозримом и отдаленном будущем.

## ПРАВДА И ВЫМЫСЕЛ

Что вам приходит на ум, когда вы слышите термин «искусственный интеллект»? Страшные роботы? Тема сложных исследований? Арнольд Шварценеггер в фильме «Терминатор»? Боты в Контр-страйке?

**i** Контр-страйк (Counter-Strike) – это компьютерная стратегическая игра-стрелялка от первого лица производства компании Valve. Она основана на битве между террористами, которые хотят устраивать взрывы бомб в людных местах, и контртеррористическими силами, мешающими террористам устроить хаос. Хотя в эту многопользовательскую игру, как правило, играют люди, предусмотрена возможность, когда один человек играет вместе с ботами или против них.

Боты – это программные виртуальные игроки с поддержкой ИИ и самостоятельным мышлением, которые могут заменять игроков, когда они отсутствуют. Боты широко используются в компьютерных играх, а иногда их просто так и называют – искусственным интеллектом игры.

Игра Counter-Strike, или CS (СиЭс), как ее с любовью называют, в особенности популярна среди любителей и профессиональных геймеров и регулярно присутствует на лучших игровых конкурсах по всему миру.

Смысл понятия «искусственный интеллект» (ИИ) изменялся в ходе исследований этого вопроса. Базовые понятия ИИ не изменились, чего не скажешь о его приложениях. То, как ИИ воспринимался в 1950-х годах, очень отличается от того, как он на самом деле используется сегодня. И понятие искусственного интеллекта по-прежнему продолжает эволюционировать.

В наши дни тема искусственного интеллекта приобрела актуальный характер. ИИ прошел долгий путь, выйдя со страниц научно-популярных фантастических романов и в результате превратившись в товар. И разумеется, ИИ не имеет ничего

общего с победным шествием роботов, которые захватывают мир и поработают нас, людей. По крайней мере, пока. Все, что обладает интеллектуальными способностями, будь то виртуальный ассистент в телефоне (Сири и Кортана), ваша верная поисковая система (Google и Bing) или ваши любимые мобильные и видеоигры, все это работает на основе ИИ.

Интерес к ИИ достиг пика в 2000-х годах, в особенности в начале 2010-х годов. Огромные инвестиции в исследования в области ИИ последних лет, направляемые научными кругами и корпорациями, стали благом для разработчиков программного обеспечения. Прогресс, достигнутый такими компаниями, как Microsoft, Google, Facebook и Amazon, в различных областях ИИ с последующим переходом к открытому исходному коду и коммерциализации своих продуктов позволил разработчикам программного обеспечения с беспрецедентной легкостью создавать впечатление человекоподобности ваших приложений. Это привело к взрывному увеличению умных, интеллектуальных приложений, которые способны понимать своих пользователей так же, как это делает обычный человек.

Задумывались ли вы, будучи разработчиком, о том, как применить ИИ для создания невероятно умного программного обеспечения? Скорее всего, задумывались, но не знали, с чего начать.

По нашему опыту общения с разработчиками программного обеспечения ведущих ИТ-компаний создается следующее впечатление: внедрение в софт даже отдельных элементов ИИ, таких как понимание естественного языка (ЕЯ), распознавание речи, машинное обучение и т. д., потребует глубокого понимания принципа работы нейронных сетей, нечеткой логики и других теоретических основ информатики.

Наверное, раньше так и было. Но поспешим вас обрадовать: таких глубоких знаний для разработки приложений с ИИ больше не потребуется. Интеллект, работающий в ваших любимых приложениях, таких как Google Search, Bing, голосовой помощник Кортана и Facebook, постепенно становится доступным для разработчиков, не работающих в этих компаниях. Некоторые составные части искусственного интеллекта этих приложений бесплатны, а другие, как SaaS (англ. software as a service – программное обеспечение как служба), предоставляются на основе коммерческих предложений.

## История и эволюция

По нашему мнению, для лучшего понимания интересующего вас предмета или вопроса, например искусственного интеллекта, следует познакомиться с его историей развития.

С древних времен люди восхищались тем, как всемогущие мудрецы либо сумасшедшие ученые наделяли неживые предметы способностью мыслить. В древней и современной литературе можно найти бесчисленное количество рассказов о том, как неодушевленные предметы внезапно наделяются сознанием и интеллектом.

Греческие, китайские и индийские философы верили, что человеческое мышление может быть представлено набором механических правил. Аристотель (384–322 до н. э.) разработал формальный метод решения силлогизмов. Евклид (~300 до н. э.) благодаря своей математической работе «Элементы», содержащей



один из самых ранних известных алгоритмов, подарил нам формальную модель рассуждения. Лейбниц (1646–1716) создал универсальный язык рассуждений, сводивший аргументацию к вычислению. Он предположил, что все рациональное мышление может быть систематизировано как алгебра или геометрия. Работа Буля (1815–1864) по математической логике была прорывной и предоставила возможность создания искусственного интеллекта.

Эти формальные системы, или «теории», с использованием технологий того времени внедрялись на практике. Создавались машины, подражавшие человеческому поведению или мыслям. Используя часовой механизм, человек создавал сложные механизмы, от часов с кукушкой до автоматов, рисующих картинку. Это были ранние варианты механических роботов. В наше время формальные принципы рассуждения были использованы для создания устройств, называемых компьютерами.

Термин «искусственный интеллект» был придуман на семинаре в кампусе Дартмутского колледжа летом 1956 года. Предлагаемая тема семинара, в частности, содержала следующее утверждение: *«Всякий аспект обучения или любое другое свойство интеллекта в принципе может быть настолько точно описано, что машина сможет его симулировать»*. Именно на этой конференции было положено начало в области исследований ИИ, а присутствовавшие на ней люди стали первоходцами научных исследований ИИ.

В последующие десятилетия произошли крупные прорывы в области разработки и изучения ИИ. Были разработаны компьютерные программы для решения алгебраических задач, доказывания теорем и ведения диалога на английском языке. Правительственные учреждения и частные организации финансировали данные исследования. Но путь развития ИИ был не прост.

Первая неудача в исследованиях ИИ произошла в 1974 году. Период между этим годом и по 1980 год называется первой «зимой ИИ». За это время многие обещанные результаты исследований не были достигнуты. Данная неудача объяснялась рядом причин, одной из которых была неспособность ученых предвидеть всю сложность проблем создания ИИ. Но основной причиной была ограниченная вычислительная мощность электронно-вычислительных машин того времени. Как результат отсутствие прогресса вынудило крупные британские и американские агентства, поддерживающие ранее эти исследования, прекратить их финансирование.

Следующие семь лет, начиная с 1980 года и по 1987 год, вновь появился интерес к исследованиям в области ИИ. Развитие экспертных систем подпитывало ажиотаж вокруг данных исследований. Экспертные системы развивались во всех организациях, и вскоре все гиганты индустрии начали вкладывать в данные исследования огромное количество денег. Исследование нейронных сетей заложило основу развития методов оптического распознавания символов и распознавания речи. В последующие годы, с 1987 по 1993 год, настала вторая зима в развитии искусственного интеллекта. Как и предыдущей зимой, финансирование разработок и исследования ИИ снова уменьшились.

**i** Экспертная система – это программа, которая отвечает на вопросы или решает задачи из конкретной области знаний, используя логические правила, которые разработаны на основе знаний экспертов.

Экспертные системы ограничивались небольшой областью конкретных знаний (при этом проблема понимания здравого смысла не возникала). Простая конструкция экспертных систем позволяла относительно легко их программировать и модифицировать после создания. В целом данные программы оказались полезными, чего ИИ ранее достичь не удавалось.

1993–2001 годы ознаменовались возвращением к исследованию и разработке искусственного интеллекта. Одной из причин возобновления интереса к ИИ стало появление более производительных и дешевых компьютеров. Закон Мура предсказывает, что скорость и объем памяти компьютеров удваиваются каждые два года. Именно это и произошло. Наконец-то прежние обещанные перспективы развития и исследования искусственного интеллекта были реализованы благодаря появлению более мощных и быстрых вычислительных машин, отсутствие которых было основной причиной неудачи еще первой зимой ИИ. Были созданы специализированные компьютеры с использованием передовых технологий ИИ, которые превосходили по своим способностям людей. Как же можно забыть культовый матч 1997 года между компьютером Deep Blue компании IBM и тогдашним действующим чемпионом мира по шахматам Гарри Каспаровым?

ИИ широко использовался в области робототехники. Взять, к примеру, японских роботов, которые выглядели как люди и даже понимали и говорили на человеческих языках. Западный мир не отставал. И вскоре началась гонка – кто первый построит наиболее человекоподобного механического помощника человека. Робот ASIMO компании Хонда был блестящим примером того, что может быть достигнуто путем объединения робототехники с ИИ: 130-сантиметровый гуманоид, который мог ходить, танцевать, готовить кофе и даже дирижировать оркестром.

## Нынешнее состояние дел

Изначально ИИ разрабатывался для создания человекоподобных роботов, которые могли бы нас понимать, выполнять нашу рутинную работу и избавлять от одиночества. Но сегодня, благодаря появлению различных методов, помогающих создавать надежные, умные и функциональные приложения, область применения ИИ стала гораздо шире.

С появлением нового поколения технологических компаний XXI век стал свидетелем огромных достижений в области искусственного интеллекта. Иногда эти достижения скрываются в тишине научно-исследовательских лабораторий Microsoft, IBM, Google, Facebook, Apple, Amazon и других. Возможно, одним из лучших примеров современного искусственного интеллекта является суперкомпьютер Watson компании IBM, первоначально разрабатывавшийся как компьютерная система, которая могла бы конкурировать с людьми. В феврале 2011 года этот суперкомпьютер принял участие в телевикторине «Jeopardy!» – аналог хорошо нам знакомой телевикторины «Своя игра». Будучи отключенным от сети Интернет, суперкомпьютер смог обыграть двух игроков: Брэда Раттера – обладателя самого большого выигрыша в программе и Кена Дженнингса – рекордсмена по длительности беспроигрышной серии. Компьютер одержал победу, получив 1 млн долларов. Дженнингс и Раттер получили, соответственно, по 300 и 200 тысяч. На волне успеха своего суперкомпьютера IBM вскоре выпустила коммерческие приложения, в которые были встроены технологии искусственного интеллекта. Интерес крупных технологических компаний к ИИ снова возрос. В результате и другие крупные технологические

компании вышли на рынок со своими собственными коммерческими предложениями. Сегодня существуют стартапы, предлагающие узкоспециализированные коммерческие предложения со встроенными технологиями ИИ.

Искусственный интеллект не был ограничен массовыми и корпоративными программными приложениями. Уже несколько лет ИИ встраивается в ваши любимые видеоигры, как телевизионные, так и мобильные. Например, в однопользовательских играх, где вы конкурируете против компьютера, ваши виртуальные противники принимают свои собственные решения в зависимости от ваших действий. Во многих играх даже можно менять уровень сложности противников: чем выше уровень сложности, тем более изощренным является искусственный интеллект игры и тем больше ваши противники становятся похожи на людей.

## Превращение ИИ в товар

За последние годы произошел взрывной, почти экспоненциальный рост объема информации. В условиях ежедневно дешевого пространства для хранения данных крупные корпорации, равно как и небольшие стартапы, стали сохранять свои архивы, чтобы в какой-то момент их проанализировать и получить важную информацию, способную принести выгоду своему предприятию. Эта тенденция в значительной части была поддержана облачной революцией. Сама облачная революция обусловлена появлением компьютеров с более высоким быстродействием и дешевым пространством хранения данных. Услуги облачного вычисления и хранения данных в облаке, предоставляемые такими компаниями, как Amazon AWS и Microsoft Azure, настолько дешевые, что уже больше не имеет смысла уничтожать даже журнальные данные десятилетней давности, произведенные серверами и программным обеспечением предприятия.

В результате каждый день и каждый час компании производят ошеломляющие объемы данных. Этот большой объем данных называется **большими данными**. Большие данные применяются почти во всех секторах экономики, таких как банковское дело, розничная торговля, ИТ, социальные сети, здравоохранение, наука, спорт и т. д.



Чтобы вообразить масштаб больших данных, обратите внимание на приведенную ниже статистику.

По состоянию на август 2017 г. поисковик Google обрабатывал примерно 100 миллиардов поисковых запросов в месяц. Это больше 1,2 триллиона запросов в год! Google анализирует свои поисковые данные с целью определения поисковых тенденций среди различных географических регионов и демографических групп.

Социальная сеть Facebook в день обрабатывает свыше 300 миллионов фотографий из своей базы пользователей. Facebook анализирует свои данные, посты и фотографии с целью предоставления своим пользователям более точных объявлений.

Крупнейшая в мире сеть оптовой и розничной торговли Wal-Mart каждый час обрабатывает свыше 1 миллиона покупательских транзакций. Wal-Mart анализирует эти данные, чтобы понимать, какие товары реализуются лучше других, какие товары реализуются вместе, и получать другую розничную аналитическую информацию.

Традиционные методы обработки данных стали нежизнеспособными из-за сложности данных и времени, которое требуется для их полного анализа. Чтобы проанализировать этот огромный объем данных, был необходим радикально новый подход. Как оказалось, с большими данными могут использоваться

и методы машинного обучения, применяемые для тренировки сложных систем ИИ. В результате сегодня ИИ больше не является прерогативой крупных частных и общественных научно-исследовательских институтов. ИИ и его разнообразная методология используются для построения и поддержки программных продуктов для всех видов компаний.

## MICROSOFT И ИИ

Компания Microsoft располагает богатой историей в области развития искусственного интеллекта. Когда в 1991 г. Билл Гейтс создал Microsoft Research, он предвидел, что однажды компьютеры будут способны видеть, слышать и понимать людей. Двадцать шесть лет спустя ИИ приблизился к воплощению таких прогнозов. За эти годы Microsoft не объявляла о создании гуманоидных роботов и не строила всезнающие универсальные компьютеры. Прогресс компании в области ИИ для общественности был «не видим». Компания потихоньку интегрировала человекоподобное мышление в свои существующие продукты.

Возьмем, к примеру, Microsoft Bing, популярную поисковую систему компании Microsoft. Мало того, что Bing может выполнять поиск по ключевым словам. Этот поисковик также может осуществлять поиск в Интернете на основе подразумеваемого смысла вашей поисковой фразы. Так, выполнение простого поиска по ключевым словам «Тейлор Свифт» дает вам официальный сайт, страницу Википедии, учетные записи в социальных сетях, последние новости и несколько фотографий популярного американского певца-автора песен. Выполнение более сложного запроса, типа «Кто президент Уганды?», даст вам точное имя крупными буквами и главные результирующие веб-страницы, относящиеся к этому человеку. Это равнозначно тому, как спросить «Кто президент Уганды» у человека, который знает, что вам нужно только имя президента данной страны, а не все веб-страницы, связанные с этим именем.

В обоих примерах (Тейлор Свифт и президент Уганды) поисковик Bing также покажет слева краткую справку о человеке: дату рождения, супругу(а), детей и т. д. И в зависимости от того, кто этот человек, вы увидите другие подробности, такие как образование, годы жизни, цитаты политика и собственный капитал, музыкальные композиции и песни для певца. Каким образом поисковой системе Bing удастся показывать вам столько много информации о человеке? Неужели разработчики этого поисковика создали мегабазу данных справочных фактов для всех мировых знаменитостей (в настоящем и прошлом)? Не совсем.

Хотя человеческих способностей хватит, чтобы создать такую базу данных, стоимость ее поддержки была бы огромной, так как наш большой и обширный мир, в котором существует множество стран и территорий, продолжает производить знаменитостей. Так что с этой базой данных возникнет определенная проблема масштабируемости.

Технология, которую компания Microsoft применила для решения этой задачи, называется *машинным обучением*. Мы рассмотрим машинное обучение и его младшего брата – глубокое обучение – чуть ниже. Поисковая система Bing научилась понимать значение поисковой фразы только благодаря подобласти ИИ, название которой – *понимание естественного языка*. Можно задать поисковику Bing тот же самый вопрос десятью различными способами, и каждый раз он по-

прежнему будет приходиться к одному и тому же смыслу. Методология понимания естественного языка (ЕЯ) достаточно умна, чтобы толковать естественно-языковые высказывания таким образом, как это люди делают подсознательно. Методология понимания ЕЯ также помогает обнаруживать в поисковых выражениях орфографические ошибки: вопрос «Кто преизидент Уганды» поисковик Bing автоматически исправит на «Кто президент Уганды».

## Фундаментальные понятия

Прежде чем вы начнете строить умные приложения с использованием элементов искусственного интеллекта, будет полезно узнать его основы. В этом разделе мы познакомимся с фундаментальной терминологией и узнаем, что происходит за кадром при работе каждого из этих терминов. Благодаря этому вы получите общее представление о том, как работает ИИ. Рисунок 1.1 показывает мимолетный взгляд в будущее, когда человек обучает машину и машина делает заметки.

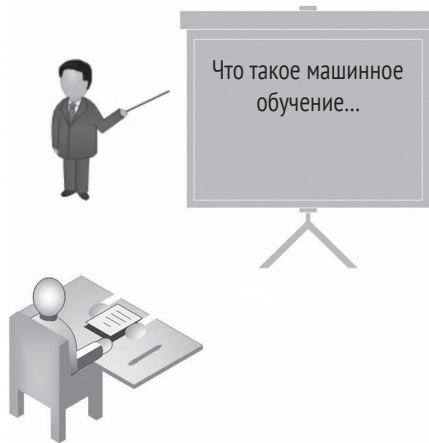


Рис. 1.1 ❖ Человек преподает машине основы, и машина делает заметки

Но, прежде чем мы сможем углубиться в изучение подробностей работы различных форм ИИ, важно уяснить одно понятие, которое приводит все эти формы в действие. Этим понятием является машинное обучение (рис. 1.1).

Термин «*машинное обучение*» был введен Артуром Сэмюэлем в своей статье 1959 года «Некоторые исследования в машинном обучении». Согласно Сэмюэлю, машинное обучение – это то, что «*придает компьютерам способность учиться, без программирования их явным образом*». Все мы знаем компьютер как машину, которая выполняет определенные операции, следуя инструкциям, поставляемым людьми в форме программ. Но тогда каким образом можно сделать так, чтобы машина училась чему-то самостоятельно? И что такая машина будет делать со знаниями, полученными в результате такого обучения?

Чтобы лучше разобраться в машинном обучении, давайте возьмем в качестве примера популярный инструмент машинного перевода Google Translate, инстру-

мент, позволяющий легко переводить фразы с иностранного языка, скажем французского на английский и наоборот. Вы когда-либо задавались вопросом: как он работает?

Рассмотрим приведенное ниже предложение на французском языке (рис. 1.2).

## Quel est ton numéro de téléphone?

Рис. 1.2 ❖ Предложение на французском языке

Надпись на рисунке: Какой ваш номер телефона?

Самая простая система перевода перевела бы это предложение пословно при помощи словаря (рис. 1.3).

## Quel est ton numéro de téléphone?

which is your number of telephone

Рис. 1.3 ❖ Буквальный перевод на английский язык

Надписи на рисунке: Который есть ваш номер телефон

Если фразу, означающую на французском языке «какой ваш номер телефона», перевести на русский язык буквально, мы получим словосочетание «который есть ваш номер телефон». Очевидно, такой упрощенный перевод полностью игнорирует правила грамматики, характерные для русского языка.

Ситуацию можно исправить, введя в систему перевода правила грамматики обоих языков. Но вот проблема: правила грамматики работают, исходя из допущения, что входное предложение грамматически оформлено правильно. В реальном мире это не всегда так. Кроме того, может иметься несколько разных правильных вариантов результирующего предложения. Такую систему, основанную на грамматических правилах перевода, стало бы слишком сложно поддерживать.

Идеальная система перевода может учиться переводить самостоятельно, получая лишь тренировочные данные. Перебрав тысячи и тысячи тренировочных фраз, система начнет видеть повторяющиеся шаблоны (образы) и в результате автоматически выводить правила языка. Машинное обучение как раз и занимается такого рода самообучением.

Google Translate поддерживает не 10 и не 20, а более 100 языков, включая некоторые редкие и малоизвестные, при этом в него регулярно вносятся дополнительные языки. Конечно же, человеческих способностей компьютеру не хватит, чтобы *жестко запрограммировать* переводы всех возможных фраз и предложений. Способность Google-переводчика понимать языки и переводить с языка на язык приводится в действие машинным обучением.

Фразы после перевода Google-переводчиком не совершенны, но их смысл сохранен. Google-переводчик учится не только на тренировочных данных, которые ему предоставляет Google, но и благодаря миллионам своих пользователей. В случае неправильного перевода пользователи могут вручную представить правильный



вариант. Google-переводчик учится на своих ошибках, точно так же, как человек, и улучшает свое понимание языков ради будущих переводов. Яркий пример машинного обучения!

**i** Совсем недавно Google-переводчик перешел с использования алгоритмов машинного обучения на алгоритмы глубокого обучения.

### **Машинное обучение против глубокого обучения**

Если вы следите за новостями, то, вероятно, вам доводилось слышать термин «глубокое обучение» в связи с искусственным интеллектом. Глубокое обучение – это новейшая разработка в области машинного обучения, и люди, которые, по-видимому, не знакомы с точным значением этого термина, понимают этот термин как преемника машинного обучения. Это совершенно неправильно.

В то время как машинное обучение является способом реализации искусственного интеллекта, глубокое обучение является методом машинного обучения. Другими словами, глубокое обучение является не альтернативой машинному обучению, а частью самого машинного обучения.

В машинном обучении общепринятой методикой традиционно были искусственные нейронные сети (ИНС). ИНС очень сильно загружают центральный процессор. Но результаты этих вычислений ниже человеческих возможностей. Недавняя революция в области ИИ стала возможной благодаря глубокому обучению, впечатляющему методу, который выполняет машинное обучение намного быстрее и точнее. В алгоритмах глубокого обучения для достижения поставленной цели (идентификации изображения, перевода с языка на язык и т. д.) применяется параллельное программирование, которое опирается на миллионы прецедентов тренировочных данных. Такое «глубокое» обучение было бы невозможным на основе предыдущих методов машинного обучения.

Компании разработали свои собственные инструменты глубокого обучения и в результате предложили облачные службы с элементами ИИ. В конце 2015 г. компания Google предоставила в свободный доступ свою платформу глубокого обучения, название которой TensorFlow. Чтобы узнать возможности этой платформы и как ее можно использовать, обратитесь по адресу: [www.tensorflow.org](http://www.tensorflow.org).

## **Машинное обучение**

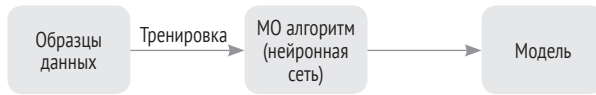
Технология машинного обучения (МО) находится в самой основе искусственного интеллекта. В рамках МО проводятся исследования и конструирование алгоритмов, которые могут обучаться на ранее обработанных данных и прогнозировать, основываясь на результатах этого обучения. Именно МО приводит в действие интеллектуальную машину, и именно на МО функционирует искусственный интеллект.

Традиционный не-МО алгоритм машинного перевода имеет статические программные инструкции, которые идентифицируют, на каком языке написано предложение: используемые слова, грамматическая структура и т. д. Не-МО алгоритм идентификации лиц аналогичным образом имеет жестко запрограммированное определение лица: нечто округлое с цветом кожи и двумя небольшими темными областями в верхней части (глазами) и т. д. С другой стороны, у МО алгоритма нет такого жесткого программирования; он учится на примерах. Если в процессе тренировки использовать большое количество предложений на французском языке и добавить небольшое количество предложений не на французском, МО алгоритм научится определять французские предложения.

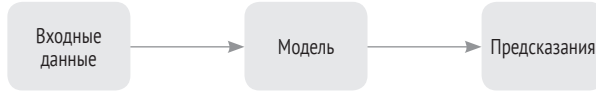
Целый ряд реальных практических задач, таких как перевод с языка на язык, предсказание погоды, фильтрация спама в почтовых сообщениях, предсказание

результатов выборов следующего президента Соединенных Штатов, задачи классификации (такие как различение видов птиц по фотографиям) и т. д., имеет нелинейный характер. МО является идеальным решением для таких нелинейных задач, где проектирование и программирование задаваемых явным образом алгоритмов с использованием статических программных инструкций попросту не работает.

Надеемся, что пример с машинным переводом в предыдущем разделе дал вам четкое понимание того, как работает машинное обучение. Это была просто верхушка айсберга. Технология МО является многосложной и многогранной. Но вы узнали, что МО – это фундаментальное понятие. МО является подобластью информатики, включающей в себя несколько тем, в частности связанных с математикой и статистикой. Хотя для охвата всех аспектов МО потребуется не одна книга, давайте взглянем на общепринятые термины, которые с ним связаны (рис. 1.4 и 1.5).



**Рис. 1.4** ❖ Алгоритм машинного обучения, в частности нейронная сеть, которая «обучается» основам темы на тренировочных данных. Результатом такого обучения является натренированная модель



**Рис. 1.5** ❖ Натренированная модель затем может принимать новые или ранее известные данные и делать информированные предсказания

Прежде чем система машинного обучения сможет начать разумно отвечать на вопросы по теме, она должна этой теме обучиться. Для этого МО в значительной степени опирается на начальный набор данных по теме обучения. Эти исходные данные называются *тренировочными данными*. Чем больше тренировочных данных, тем больше шаблонов (образов) наша машина в состоянии идентифицировать и тем точнее она сможет отвечать на вопросы – новые и ранее известные – по этой теме. Для получения надежных результатов обычно недостаточно нескольких сотен или даже тысяч записей тренировочных данных.

Действительно точные, человекоподобные машины были натренированы с использованием миллионов записей или нескольких гигабайтов данных в течение дней, месяцев или даже лет. И мы ни на йоту не преувеличиваем. Продолжительность непрерывной работы персонального компьютера с хорошей вычислительной мощностью и высококачественной видеокартой для тренировки алгоритма машинного перевода данными объемом свыше 1 Гб только для одной пары языков составит больше месяца [см. <https://github.com/tensorflow/tensorflow/issues/600#issuecomment-22633266>].

Качество тренировочных данных и то, как модель сконструирована, имеют равноценное значение. Используемые данные должны быть точными, очищен-



ными и поступать через надежные источники. Модель должна быть разработана с использованием сценариев реальной практической деятельности. Так что в следующий раз, когда ваше приложение распознавания образов неправильно идентифицирует захваченный объект или ваше любимое приложение машинного перевода выдаст смешной перевод, вините в этом качество тренировочных данных или модель, которая использовалась. Кроме того, важно отметить, что обучение не является процедурой начального этапа: это непрерывный процесс. Первоначально машина обучается на тренировочных данных; позже она делает это на пользователях.

Исследование в области ИИ привело к развитию нескольких подходов к реализации машинного обучения. Искусственная нейронная сеть является одним из самых популярных подходов. ИНС, или просто *нейросеть*, является обучаемым алгоритмом. Основой для разработки нейросети послужили структура и функциональные аспекты биологических нейронных сетей. Вычисления структурированы с точки зрения взаимосвязанных групп искусственных нейронов, обрабатывающих информацию, используя ассоциативный подход к компьютерным вычислениям. Они используются для моделирования сложных связей между входами и выходами с целью отыскания этих связей в данных шаблонов (образов). Другие популярные подходы представлены *глубоким обучением*, системами на основе *продукционных правил* (если – то), *разветвленными алгоритмами решений* и *байесовскими сетями*.

Таким образом, когда в нейронные сети подается достаточный объем тренировочных данных, мы получаем то, что называется натренированной моделью. *Модели* – это математические и статистические функции, которые могут создавать *прогнозы* (информированные предположения) в отношении предоставленных входных данных. Например, основываясь на метеорологической информации за прошедшие 10 лет (тренировочных данных), модель машинного обучения может научиться предсказывать погоду на несколько дней вперед.

### **Типы машинного обучения**

*Обучение с учителем*, или контролируемое обучение, выполняется, когда тренировочные данные помечены. Если в процессе обучения алгоритма распознавания языка предложения, передаваемые в этот алгоритм, будут помечены, то такое обучение называется обучением с учителем. То есть в этом случае требуется передавать алгоритму предложения, явным образом помеченные языком, на котором они написаны. Например, предложение на французском и не на французском языке, на испанском и не на испанском языке и т. д. Поскольку предварительная разметка выполняется людьми, эта работа увеличивает трудозатраты и стоимость поддержания таких алгоритмов.

*Обучение без учителя*, или неконтролируемое обучение, выполняется, когда тренировочные данные не помечены. Из-за отсутствия меток, разумеется, алгоритм не может волшебным образом научиться правильно определять язык предложения, но он может дифференцировать один язык от другого. Таким образом, используя обучение без учителя, алгоритм МО может научиться определять, что французские предложения отличаются от испанских, которые отличаются от хинди, и т. д.

*Обучение с подкреплением*, или стимулированное обучение, выполняется, когда машине не передаются тренировочные данные явным образом. Чтобы достигнуть

поставленной цели, алгоритм должен взаимодействовать с окружающей средой. Из-за отсутствия тренировочных данных он должен обучаться самостоятельно с нуля и полагаться на метод проб и ошибок, чтобы принимать решения и обнаруживать свои собственные правильные пути. У каждого выполняемого машиной действия есть последствие, и каждому последствию задано числовое вознаграждение. Таким образом, если действие приводит к желательному результату, оно получает «хорошие» отметки. И если результат имеет катастрофические последствия, то он получает отметки «очень и очень плохо». Как и люди, машина стремится максимизировать свое итоговое числовое вознаграждение – то есть получить как можно больше «хороших» и «очень хороших» отметок, не повторяя свои ошибки. Этот метод машинного обучения особенно полезен, если машина будет работать в быстро изменяющихся средах, где создание и поставка тренировочных данных попросту невозможны. Это такие среды, как управление автомобилем (рис. 1.6), видеоигры и т. д.



**Рис. 1.6** ❖ Самоуправляемые автомобили – автотранспортные средства, которые не требуют, чтобы ими управлял человек, используют обучение с подкреплением, чтобы обучаться на данных динамической и сложной среды (дороги и дорожное движение) с целью постепенного улучшения своих навыков вождения

## Язык

Люди взаимодействуют друг с другом одним из трех способов: словесно, письменно и жестами. Эти три способа объединены в одно понятие – «язык». *Язык* – это свод правил общения, который одинаков для каждого человека. Общение может быть устным и письменным, и между этими двумя способами общения есть различия. Формальной формой общения является письмо. При этом язык знаков, язык жестов полностью отличается.

Большая часть усилий, приложенных в исследовании ИИ, состояла в том, чтобы наделить машины способностью понимать людей так же естественно, как это делают сами люди. Поскольку машинам легче понимать письменную речь, чем устную, мы начнем наше обсуждение с основ понимания языка в письменной форме.

### ***Понимание естественного языка***

*Понимание естественного языка* (ЕЯ) – это способность машины понимать людей посредством естественных языков. Компьютер изначально разработан, чтобы по-

нимать биты и байты, программный код и логику, программы и инструкции, а не естественные языки. Иными словами, компьютер умело управляет структурированными данными, чего нельзя сказать о неструктурированных данных.

Естественный язык регламентируется некими правилами (грамматикой), однако во время ежедневного и неофициального общения эти правила не всегда соблюдаются. В результате люди легко могут понимать письменные или словесные высказывания с плохой грамматикой, неправильным произношением, просторечиями, сокращениями и т. д. Можно смело сказать, что естественные языки регламентируются гибкими правилами.

Технология понимания ЕЯ превращает неструктурированные исходные данные (рис. 1.7), регламентированные гибкими и плохо определенными правилами, в структурированные данные, которые машина способна понять. Для любознательных отметим, что именно эта технология делает человекоподобными Microsoft Кортана, Apple Сири и Amazon Алекса.

## Какая завтра ожидается погода в Дели

| День   | Намерение | Город |
|--------|-----------|-------|
| Завтра | Погода    | Дели  |

**Рис. 1.7** ❖ Технология понимания ЕЯ анализирует каждое предложение с целью получения двух вещей: намерения (смысла, значения или предполагаемого действия) и сущностей. В этом примере извлеченная информация о *погоде* является распознанным намерением, а *город* (Дели) и *день* (завтра) являются сущностями. Пользователь может задать тот же самый вопрос сотнями различных способов, и хорошая система понимания ЕЯ всегда будет в состоянии извлечь из пользовательского запроса правильное намерение и сущности.

Далее программная система сможет использовать эту ранее извлеченную информацию для подготовки запроса к онлайн-метеорологическому API и показать пользователю требующуюся информацию о погоде

### Обработка естественного языка

Разумеется, в человеко-компьютерном взаимодействии гораздо больше задач, чем только одно понимание намерения заданного предложения. Технология обработки ЕЯ охватывает все, что имеет отношение к человеко-компьютерному взаимодействию на естественном языке. И технология понимания ЕЯ является всего лишь подмножеством задач в большем множестве, которым является технология обработки ЕЯ. Технология обработки естественного языка, в частности, включает в себя такие задачи, как:

- **машинный перевод:** преобразование текста из одного языка в другой;
- **генерация естественного языка:** обратная по отношению к пониманию ЕЯ задача; преобразование структурированных данных (обычно из баз данных) в текстовые предложения, понятные человеку. Например, на основе сравнения двух строк в базе данных с информацией о погоде может быть сформулировано вот такое предложение: «сегодня в Дели температура воздуха составляет 26 градусов по Цельсию, что ниже на 2 градуса, по сравнению со вчерашним днем»;
- **анализ мнений:** просканировать фрагмент текста (твит, обновление статуса Facebook, обзоры, комментарии и т. д.), касающийся товара, человека, события или географического места, чтобы определить общее мнение (отрицательное или положительное) в отношении соответствующего субъекта;

- **распознавание именованных сущностей:** в том или ином тексте определить, какие фрагменты в тексте соотносятся с именами собственными, связанными с людьми или географическими объектами, и тип каждого определенного имени (человек, местоположение, организация);
- **извлечение связей:** извлечение связей между сущностями, присутствующими во фрагменте текста, в частности родственные отношения, причины, признаки и т. д.

Обработка ЕЯ намного шире, чем несколько упомянутых выше задач, причем каждая из них является предметом независимого исследования.

## Речь

Помимо интеллектуального анализа текста, ИИ способен помочь машинам с акустическими устройствами, такими как микрофон, понимать, о чем говорится. Устная речь представляется как набор звуковых сигналов, и акустическое моделирование используется для нахождения связей между звуковым сигналом и фонемами (минимальными смыслоразличительными единицами языка, которые составляют устную речь).

### *Распознавание речи*

*Распознавание речи* – это распознавание и преобразование компьютерами устного языка в текст. Когда вы задаете вопрос Siri или Google (голосовой поиск), они используют методологию распознавания речи для преобразования вашего голоса в текст. Преобразованный текст затем используется для выполнения поискового запроса. Современные методы распознавания речи умеют справляться с отклонениями в акцентах, похоже звучащими словами и фразами на основе контекста.

Применения распознавания речи колеблются от проектирования общедоступных систем (таких как программное обеспечение для слепых) или систем на основе голосового поиска до голосового (бесконтактного) набора текста.

### *Распознавание голоса*

Термины «*распознавание голоса*» и «*идентификация говорящего*» относятся к идентификации диктора, а не к смыслу произносимой фразы. Распознавание говорящего может упростить задачу перевода речи в системах, натренированных на голос определенного человека. Распознавание голоса диктора может использоваться с целью подтверждения подлинности или проверки личности говорящего и служить в качестве составной части процесса обеспечения безопасности.

### *Преобразование текста в речь и речи в текст*

Преобразования текста в речь (текст-в-речь, Text-to-Speech, TTS) и речи в текст (речь-в-текст, Speech-to-Text, STT) являются взаимосвязанными, но разными технологиями.

Преобразование текста в речь, также именуемое *синтезом речи*, – это способность машины «проговаривать» фрагмент письменного текста. Синтезируемая речь создается путем соединения частей записанной речи (записей каждого слова), которые хранятся в базе данных. В альтернативном варианте синтезатор может содержать модель речевого тракта и других характеристик человеческого голоса с целью создания абсолютно «синтетического» голосового результата на выходе.

С другой стороны, преобразование устной речи в текст является следующим шагом в распознавании речи. После того как речь была разложена на звуковые сигналы и затем на фонемы, компьютер может преобразовывать эти фонемы в текст. Используя тот же самый набор и последовательность фонем, можно построить многочисленные текстовые предложения, в результате чего машина разумно назначает каждой конструкции степень уверенности для придания предложениям большей осмысленности.

## Компьютерное зрение

Наконец, мы дошли до раздела, где обсудим методы искусственного интеллекта, с помощью которых происходит обработка визуальных данных. Визуальные данные включают в себя как неподвижные изображения – фотографии, так и динамически изменяющееся изображение – видео. Общее название этого термина – компьютерное зрение, т. е. способность компьютера «видеть». Компьютеры по своей природе не способны обрабатывать изображения так же хорошо, как текст. Методы обработки изображений в соединении с интеллектуальными алгоритмами ИИ позволяют машинам «видеть» изображения, идентифицировать и распознавать объекты и людей.

### Обнаружение объектов

Сцена на фотографии может содержать десятки или даже сотни объектов. Но обычно в зависимости от ситуации внимание фокусируется только на нескольких объектах. Давайте назовем их «интересующими». Распознавание изображения обозначает способность машины обнаруживать интересные объекты, которые могут меняться в зависимости от ситуации. Например:

- ускоряющийся автомобиль на дороге (управление дорожным движением) (рис. 1.8);



Рис. 1.8 ❖ Автомобиль, обнаруживаемый на дороге

- планетоподобный объект в обширной Солнечной системе или галактике (астрономия);
- грабитель, проникающий через задний двор (безопасность жилища);
- группа людей, входящая в торговый центр (ведение учета посещаемости).



### Распознавание образов

За обнаружением обычно следует распознавание. Распознавание – это способность распознать и назначить обнаруженным объектам и действиям точный тип (рис. 1.9). Например:

- распознавание лодки, двух человек, воды и солнца на изображении;
- распознавание точных видов животных на фотографии.

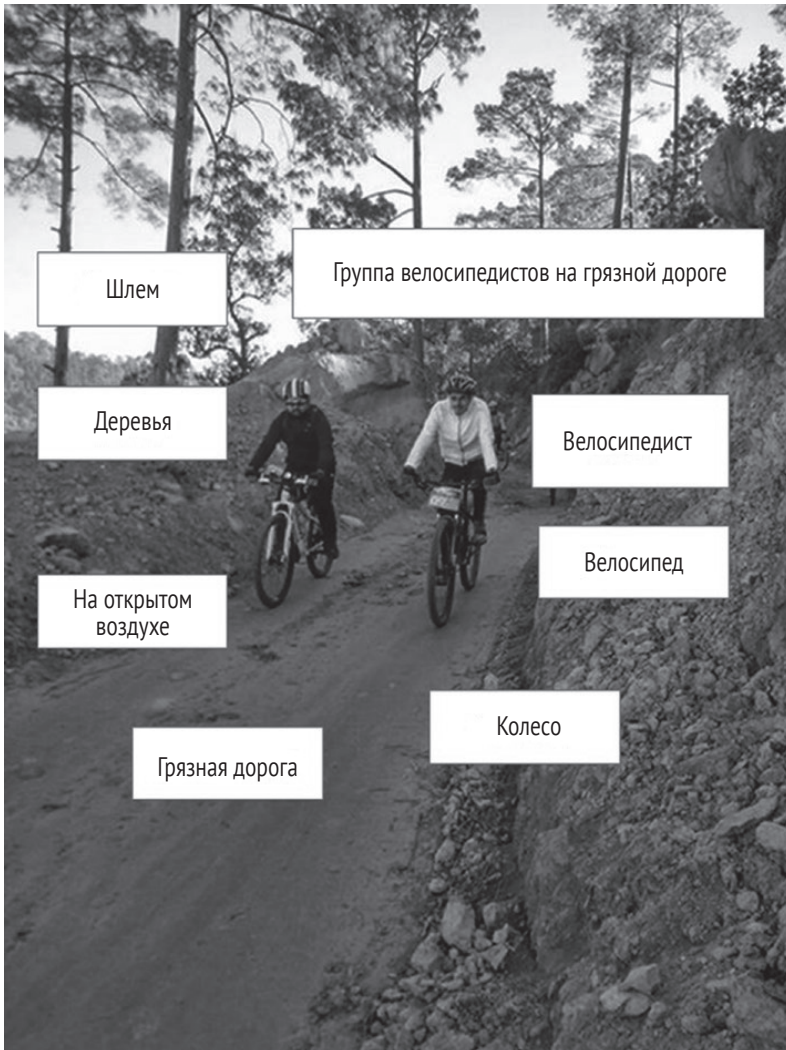


Рис. 1.9 ❖ Группа велосипедистов на грязной дороге

Распознавание образов также называется *классификацией*, или *сравнением*, объектов. Среди других систем распознавание объектов характерно для приложений виртуальной реальности, таких как Google Goggles, предназначенное для визуального поиска и распознавания изображений.

Точность системы распознавания образов, как и все остальное в ИИ, в значительной степени зависит от тренировочных данных. При помощи методов машинного обучения, чтобы научиться распознавать объекты определенного класса, система проходит тренировку на сотнях изображений. Так, можно сначала натренировать систему распознавать обычную собаку, используя сотни изображений с одной или несколькими собаками. После того как система в состоянии распознавать собак, затем ее можно натренировать распознавать немецкую овчарку, добермана или даже чихуа-хуа.

### **Распознавание лиц**

Обнаружение и распознавание лиц являются одной из многих подзадач распознавания образов (рис. 1.10). С помощью этих методов система способна обнаруживать на фотографии не только лица, но и связанные с ними признаки (возраст, пол, улыбку и т. д.). И если система предварительно была натренирована на лицо определенного человека, она может выполнить сравнение, чтобы распознать лицо именно этого человека на фотографии.

Распознавание лиц при использовании камер скрытого видеонаблюдения может использоваться в качестве механизма проверки подлинности или для обнаружения опасного преступника в местах массового скопления людей.



Рис. 1.10 ❖ Лица, определяемые по изображению

### **Оптическое распознавание символов**

Оптическое распознавание символов (ОПС, Optical Character Recognition, OCR) – это метод, используемый для преобразования рукописного, набранного на печатной машинке или распечатанного на бумаге текста в данные, которые можно отредактировать на компьютере. Система оптического распознавания просматривает сканированные изображения печатных документов и сравнивает формы отсканированных букв с эталонными изображениями букв. В результате этого создается текстовый файл, который можно отредактировать в обычном текстовом редакторе. Текст, обнаруженный с использованием ОПС, затем можно передать в систему преобразования текста в речь (TTS), чтобы проговорить его громко вслух слепому человеку, который в противном случае не смог бы этот документ увидеть.

Технология OCR обычно используется онлайн-магазинами для создания электронных копий печатных книг. Она также используется некоторыми инструментами машинного перевода, чтобы переводить вывески с иностранного языка, используя камеру мобильного телефона. Рисунок 1.11 показывает, как Google применяет OCR, превращая ваш телефон в переводчика в реальном режиме времени.



**Рис. 1.11** ❖ Google превращает ваш телефон в переводчика, действующего в реальном режиме времени

## КОГНИТИВНЫЕ СЛУЖБЫ MICROSOFT

*Когнитивные службы* (Cognitive Services) – это набор программного обеспечения Microsoft, предлагаемого на коммерческой основе в виде служб SaaS со встроенным искусственным интеллектом. Когнитивные службы представляют собой результат многолетних исследований компании Microsoft в области когнитивных вычислений и искусственного интеллекта, и многие из этих служб применяются в некоторых собственных популярных продуктах Microsoft, таких как Bing (поиск, карты), Переводчик, среда для создания программ-роботов Bot Framework и т. д.

Компания Microsoft предоставляет эти службы в виде простых в использовании веб-служб на основе механизма REST API, работающих непосредственно в веб-приложении либо мобильном приложении. На момент написания этой книги в наличии имелось 29 когнитивных служб, тематически подразделенных на пять категорий (табл. 1.1).

**Таблица 1.1. Когнитивные службы Microsoft**

|   |  |
|---|--|
| <p><b>Оптические</b></p> <ul style="list-style-type: none"> <li>• API компьютерного зрения</li> <li>• API модерации информационного наполнения</li> <li>• API анализа эмоций</li> <li>• API распознавания лиц</li> <li>• API анализа видео</li> <li>• Пользовательская служба компьютерного зрения</li> <li>• Индексатор видео</li> </ul> | <p><b>Речевые</b></p> <ul style="list-style-type: none"> <li>• API Bing обработки речи</li> <li>• Пользовательская служба обработки речи</li> <li>• API распознавания говорящего</li> <li>• API машинного перевода речи</li> </ul> |
|---|--|



Окончание табл. 1.1

|   |  |
|---|--|
| <b>Язык</b> <ul style="list-style-type: none"> <li>• API Bing проверки орфографии</li> <li>• Интеллектуальная служба понимания языка</li> <li>• API лингвистического анализа</li> <li>• API анализа текста</li> <li>• API машинного перевода</li> <li>• API веб-модели языка</li> </ul> | <b>Знание</b> <ul style="list-style-type: none"> <li>• API академических знаний</li> <li>• Интеллектуальная служба связывания именованных сущностей</li> <li>• API исследования знаний</li> <li>• API создания вопросно-ответных систем</li> <li>• API рекомендаций</li> <li>• Пользовательская служба принятия решений</li> </ul> |
| <b>Поиск</b> <ul style="list-style-type: none"> <li>• API Bing автозаполнения</li> <li>• API Bing поиска изображений</li> <li>• API Bing поиска новостей</li> <li>• API Bing поиска видео</li> <li>• API Bing веб-поиска</li> <li>• API Bing пользовательского поиска</li> </ul>        |  |

## Зрение

Службы компьютерного зрения (Vision) работают с визуальной информацией, главным образом в форме изображений и видео.

- **API компьютерного зрения.** Извлекает важную информацию из изображения, касающуюся его содержимого: интеллектуальное текстовое описание изображения, обнаруженные лица (с возрастом и полом), доминирующие на изображении цвета, и есть ли на изображении контент для взрослых.
- **Модерация информационного наполнения.** Оценивает текст, изображения и видео на наличие оскорбительного и нежелательного содержания.
- **API распознавания эмоций.** Анализирует лица с целью обнаружения спектра чувств, таких как гнев, счастье, печаль, страх, удивление и т. д.
- **API распознавания лиц.** Обнаруживает человеческие лица и сравнивает похожие лица (распознавание лиц), организует людей в группы согласно визуальному подобию (группирование лиц) и опознает ранее помеченных людей по изображениям (верификация лиц).
- **API распознавания видео.** Интеллектуальная обработка видеоданных для распознавания лиц, обнаружения движения (широко используется в системах скрытого видеонаблюдения), генерирование миниатюр изображений и анализ видео в режиме почти реального времени (текстовое описание каждого кадра).
- **Индивидуализированные (пользовательские) службы компьютерного зрения.** Когда требуется выполнить распознавание образов применительно к вещам, которые не относятся к фону, лицам и эмоциям, эти службы позволяют создавать индивидуальные классификаторы изображений, обычно ориентированные на конкретную область применения. Эту службу можно натренировать – скажем, определять различные виды птиц – и затем использовать ее REST API (репрезентативная передача состояния) в мобильном приложении для любителей наблюдения за птицами.
- **Индексатор видео.** Извлекает полезную информацию из видео с использованием таких инструментов, как распознавание лиц (имена людей), речевой анализ мнений (положительный, отрицательный, нейтральный) по каждому человеку и извлечение ключевых слов.

## Речь

Эти службы работают с человеческой речью в форме аудиоформата.

- **API Bing обработки речи.** Преобразует речь в текст, извлекает смысл высказывания и преобразовывает текст назад в речь. Этот API подробно рассмотрен в главе 7.
- **Служба персонализированной обработки речи.** Позволяет создавать индивидуальные модели языка для распознавания речи путем его увязки со словарем приложения и разговорным стилем ваших пользователей. Этот API подробно рассмотрен в главе 7.
- **API распознавания говорящего.** Опознает говорящего в записанном или живом речевом аудио. Распознавание говорящего может надежно использоваться в качестве механизма проверки подлинности.
- **API машинного перевода речи.** Переводит речь с одного языка на другой в режиме реального времени между девятью поддерживаемыми языками.

## Язык

Эти службы работают с пониманием естественного языка, переводом, анализом и многим другим.

- **API Bing проверки орфографии.** Исправляет орфографические ошибки в предложениях. Помимо словаря, принимает во внимание разрывы между словами, сленг (жаргонизмы), имена людей и названия фирменных брендов. Этот API подробно рассмотрен в главе 5.
- **Интеллектуальная служба понимания языка (Language Understanding Intelligent Service, LUIS).** Служба понимания естественного языка (ЕЯ). Этот API подробно рассмотрен в главах 4 и 6.
- **API лингвистического анализа.** Выполняет разбор текста с использованием детального лингвистического анализа, такого как выделение предложений и их лексемизация (разбиение текста на предложения и предложений на лексемы) и частеречная разметка (назначение меток существительного, глагола и т. д.). Этот API подробно рассмотрен в главе 6.
- **API анализа текста.** Обнаруживает в тексте мнение (положительное или отрицательное), ключевые фразы, темы и язык. Этот API подробно рассмотрен в главе 5.
- **API машинного перевода.** Переводит текст с одного языка на другой и идентифицирует язык данного текста. Этот API подробно рассмотрен в главе 5.
- **API веб-модели языка.** Обеспечивает множество задач обработки естественного языка, не охваченных другими языковыми API: разбиение на слова (вставка пробелов в последовательность слов с недостающими пробелами), совместные вероятности (вычисление частоты совместного появления конкретной последовательности слов), условные вероятности (вычисление частоты следования конкретного слова за другим) и варианты следующего слова (получение списка слов, которые, скорее всего, должны появиться следующими). Этот API подробно рассмотрен в главе 5.

## Знание

Эти службы работают с поиском по большим базам данных с целью идентификации сущностей, предоставления результатов поиска и рекомендаций в отношении продукта.

- **API академических знаний.** Позволяет извлекать информацию из академического графика, Microsoft Academic Graph, собственной базы знаний научных/академических научно-исследовательских работ и их содержание. Используя этот API, можно легко найти статьи по автору, научному институту, событию и т. д. Также есть возможность находить аналогичные работы, выполнять проверку на наличие плагиата и получать статистику цитирования.
- **Интеллектуальная служба связывания именованных сущностей.** Находит в тексте ключевые слова (именованные сущности, события, местоположения и т. д.) на основе контекста.
- **Служба исследования знаний.** Добавляет поддержку запросов на естественном языке, предлагает варианты поиска с автозавершением и другие возможности по отношению к вашим собственным данным.
- **Генератор вопросов и ответов.** Чудесным образом создает из предоставленных данных вопросы и ответы в стиле часто задаваемых вопросов. Создатель вопросов и ответов (QnA maker) предлагает совместную работу веб-сайта и API. Использует веб-сайт для создания базы знаний, применяя ваш существующий веб-сайт, PDF, файл Word или обычный текстовый файл часто задаваемых вами вопросов. Генератор вопросов и ответов автоматически извлекает вопросы и ответы из вашего документа (документов) и, основываясь на ваших данных, тренируется давать ответы на вопросы пользователей на естественном языке. Его можно рассматривать как автоматизированную версию LUIS. Вам не приходится тренировать систему, но при этом вы имеете возможность переобучать ее согласно вашим индивидуальным требованиям. API генератора вопросов и ответов представляет собой конечную точку, которая принимает пользовательские запросы и отправляет ответы для вашей базы знаний. При необходимости генератор вопросов и ответов может быть соединен с инфраструктурой для создания программ-роботов Microsoft Bot Framework с целью разработки готовых к работе ботов для Facebook, Skype, Slack и других каналов связи.
- **API рекомендаций.** Этот инструмент в особенности полезен для розничных, онлайн-овых и обычных магазинов. Данный API помогает увеличивать продажи магазинам, предлагая своим клиентам рекомендации, такие, например, как часто покупаемые наборы товаров. Предлагает рекомендации по товарам, которые могут заинтересовать конкретного покупателя. Эти рекомендации выбираются исходя из истории транзакций товара и т. д. Подобно генератору вопросов и ответов, веб-сайт рекомендаций в своей системе использует ваши существующие данные для создания продуктового каталога и данных об использовании.
- **Служба персонализированного принятия решений.** Использует заданную текстовую информацию для получения контекста, на основе которого предоставленные варианты можно ранжировать по приоритетности и при-

нимать решение на основе полученного рейтинга. Данная служба использует машинное обучение с подкреплением на основе обратной связи с целью постепенного совершенствования своей работы.

## Поиск

Эти службы помогают задействовать возможности поисковой системы Bing, занимающей по популярности второе место в мире.

- **API Bing автозаполнения.** Предоставляет вашему приложению форму поиска и, когда пользователь набирает запрос в поле поиска, параллельно с этим предоставляет интеллектуальный опережающий ввод запроса и варианты результатов поиска непосредственно в Bing.
- **API Bing поиска изображений.** Использует поиск изображений Bing и предоставляет изображения на основе фильтров, таких как ключевые слова, цвет, страна, размер, лицензия и т. д.
- **API Bing поиска новостей.** Выводит результаты поиска свежих новостей на основе фильтров, таких как ключевые слова, новизна, страна и т. д.
- **API Bing поиска видео.** Выполняет поиск видео на основе фильтров, таких как ключевые слова, разрешение, продолжительность видео, страна и цена (бесплатные или платные).
- **API Bing поиска в вебе.** Производит поиск в вебе на основе различных фильтров. Также есть возможность получать список похожих запросов для ключевого слова или фразы.
- **API Bing индивидуализированного (пользовательского) поиска.** Назначение этой службы – поиск в Bing на основе индивидуальных значений и тем. В результате вместо поиска по всему вебу поисковик Bing будет искать веб-сайты, основываясь на теме (темах). Он может также использоваться для реализации локализованного поиска по одному или нескольким веб-сайтам из заданного списка.

Все API Когнитивных служб имеют две ценовые категории: бесплатную и с оплатой по мере использования. Ценовую категорию можно выбрать, исходя из активности эксплуатации вашего приложения.

К сожалению, описать каждый из этих API в мельчайших подробностях не позволяет объем данной книги. Мы рассмотрим несколько служб из категорий речи, языка и поиска, которые смогут обеспечить вас стартовыми условиями для создания в короткие сроки действительно умных приложений.

Подробнее узнать об этих службах и обо всех внесенных изменениях можно, обратившись по адресу [www.microsoft.com/cognitive-services/en-us/apis](http://www.microsoft.com/cognitive-services/en-us/apis).

## РЕЗЮМЕ

Эта глава послужила введением в технологию искусственного интеллекта. Вы познакомились с историей, методами и базовой терминологией ИИ. Мы рассказали об исследованиях, проводимых компанией Microsoft в области искусственного интеллекта. Также получили краткий обзор различных коммерческих предложений от компании Microsoft, в которых применены элементы ИИ в форме когнитивных служб, основанных на механизме REST API.

Подводя итоги, в этой главе вы:

- узнали правду и вымысел об ИИ;
- познакомились с историей и эволюцией искусственного интеллекта;
- узнали, как и где ИИ используется сегодня;
- познакомились с машинным обучением, которое в действительности является основой любой интеллектуальной системы;
- познакомились с готовыми к применению Когнитивными службами компании Microsoft, основанными на программных интерфейсах REST API и использующимися для создания интеллектуальных приложений.

В **следующей главе** вы узнаете, какие предварительные требования необходимо соблюсти и какие инструменты установить для создания программного обеспечения с поддержкой ИИ, после чего создадите свое первое умное приложение с использованием программной среды Visual Studio.