SAGE ESSENTIALS

AI and machine learning

Was Rahman

SAGE | Response Business Books

SAGE was founded in 1965 by Sara Miller McCune to support the dissemination of usable knowledge by publishing innovative and high-quality research and teaching content. Today, we publish over 900 journals, including those of more than 400 learned societies, more than 800 new books per year, and a growing range of library products including archives, data, case studies, reports, and video. SAGE remains majority-owned by our founder, and after Sara's lifetime will become owned by a charitable trust that secures our continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne

Advance Praise

Artificial intelligence and machine learning are key components and driving forces behind the Fourth Industrial Revolution that is influencing and changing our personal and professional lives. This book is an excellent source of questions and answers that are comprehensible and usable by people at all levels of knowledge, interest and involvement spectrum.

—**Larry Hirst CBE,** Former Chairman, IBM EMEA;
Member and Former Chairman of Imperial College
Data Science Institute Board

Was Rahman's *AI and Machine Learning* is a must-read! It will appeal to a wide range of readers, including my philosopher Dad and my psychology student niece, and has a great deal of relevance and application to my work in sports and media. Rahman takes us on a fascinating tour of the history of AI, key terms and concepts; how AI and ML are used today, and ends with a futurist look at what may happen next. An important and accessible book at a critical time for humanity and thinking machines!

—**Anna Lockwood,**
Head of Global Sales, Telstra Broadcast Services

Mr Rahman's book is a great summary of complex topics related to artificial intelligence and machine learning. The author has done extensive research on these topics and has shared it with us in a very easy-to-read and digestible manner with excellent tie-ins to real-life and practical examples.

—Antony Stephen,
General Manager, EU Payment Products, Amazon

If you are looking for an introduction to artificial intelligence, this is the book—rigorous but not mathematical, simple yet profound. I have no hesitation in saying that it is among the best books I have read in recent years!

—Professor C. T. Kurien (retired),
Former Director and Chairman,
Madras Institute of Development Studies;
Former National Fellow of the
Indian Council of Social Science Research

It is a simple and carefully written explanation of the concepts underlying AI and the history of their development. It is a good read for technology enthusiasts seeking to explore the primary concepts of what makes AI and ML valuable while simultaneously examining the potential risks of AI misuse.

—Miguel de Andrés-Clavera,
Senior Product Manager, Google

AI and machine learning

S A G E   E S S E N T I A L S

The *SAGE Essentials* series aims to arm professionals with bite-sized learning on technical and relevant topics, helping them to stay ahead in the knowledge-powered economy. It offers a range of pithy books on topics of current interest for business and management professionals. Books in the series, authored by subject matter experts, provide a historically entrenched preamble to the subject, following it up with invaluable insights, practical ideas and future trends. In a volatile and rapidly changing business and technological environment, these short and succinct books aim to help professionals from diverse backgrounds and industries upskill by broadening their knowledge and enhancing their learning capacities.

AI and machine learning

Was Rahman

SAGE | Response Business Books

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

SAGE ESSENTIALS

First published in 2020 by

SAGE Publications India Pvt Ltd
B1/I-1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044, India
www.sagepub.in

SAGE Publications Inc
2455 Teller Road
Thousand Oaks, California 91320, USA

SAGE Publications Ltd
1 Oliver's Yard, 55 City Road
London EC1Y 1SP, United Kingdom

SAGE Publications Asia-Pacific Pte Ltd
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Published by Vivek Mehra for SAGE Publications India Pvt Ltd. Typeset in 9/13.5 pt Georgia by AG Infographics, Delhi.

Library of Congress Cataloging-in-Publication Data

Names: Rahman, Was, 1965- author.
Title: AI and machine learning / Was Rahman.
Description: New Delhi, India : SAGE Response, an imprint of SAGE Publishing, 2020. | Includes bibliographical references.
Identifiers: LCCN 2020024955 | ISBN 9789353884413 (paperback) | ISBN 9789353884420 (epub) | ISBN 9789353884437 (ebook)
Subjects: LCSH: Artificial intelligence. | Machine learning.
Classification: LCC Q335 .R34 2020 | DDC 006.3—dc23
LC record available at https://lccn.loc.gov/2020024955

ISBN: 978-93-5388-441-3 (PB)

SAGE Team: Namarita Kathait, Shruti Gupta, Parul Prasad and Kanika Mathur

Dedication

To Zayn and Gabriel,
who'll have to live with the consequences
of what we're doing to their world.

Thank you for choosing a SAGE product!
If you have any comment, observation or feedback,
I would like to personally hear from you.

Please write to me at **contactceo@sagepub.in**

Vivek Mehra, Managing Director and CEO, SAGE India.

Bulk Sales

SAGE India offers special discounts
for purchase of books in bulk.
We also make available special imprints
and excerpts from our books on demand.

For orders and enquiries, write to us at

Marketing Department
SAGE Publications India Pvt Ltd
B1/I-1, Mohan Cooperative Industrial Area
Mathura Road, Post Bag 7
New Delhi 110044, India

E-mail us at **marketing@sagepub.in**

Subscribe to our mailing list

Write to **marketing@sagepub.in**

This book is also available as an e-book.

Contents

SAGE ESSENTIALS

Foreword

Was Rahman's 'SAGE Essentials', about artificial intelligence (AI) and machine learning (ML) is a crisp exploration of the ways these powerful trends will shape our daily lives for decades to come. It weaves a compelling story from the early history to the current day and provides a vision for future uses. Rahman efficiently articulates what these trends mean in the human and computer world, crafting a light, easy read that satiates the curious mind without having to learn Python coding, TensorFlow or MXNet.

Rahman's examination of AI history is a rich read, detailing the ways philosophers, mathematicians, scientists and engineers have enabled its current form. The author ably connects the dots, noting the scores of early projects, from the British Science Research Council to progress made by Defense Advanced Research Projects Agency (DARPA), Dartmouth College and the Japanese government. He reminds us of Hollywood fanning the imaginative flames for J.A.R.V.I.S (Just A Rather Very Intelligent System) in Iron Man, and how it made us fear the dark possibilities of HAL in *2001: A Space Odyssey*.

Rahman's six key concepts for AI and ML are well defined and highly applicable to a variety of uses, from computer vision to natural language processing and process automation. Readers will identify with the 'joy' of deep learning, even when it springs from teaching a machine.

This is a necessary read for the curious-minded business leaders, scientists, doctors, educators and students. Each will find value here; from the corporation seeking an analysis of AI's strengths, weaknesses, opportunities and threats, to the student considering myriad fields of study or career. Rahman walks through common applications of daily use—smart speakers, home automation, navigation systems, shopping-prediction algorithms, insurance quotes and logistics—and others, such as human resources where AI application is legally complex, to give readers a broader understanding of how, when and where AI can be used.

There is a plethora of knowns, and unknown unknowns, in AI. And while many fear this future, Rahman goes to the core questions, including control, morality and legality. Corporate AI leaders such as Amazon, Facebook, Alibaba and Google are all trying to develop a firm grasp of these issues.

Besides presenting his own views on the future of AI, Rahman assimilates views from some of the great thinkers on the topic: Kurzweil, Asimov, Kaku, Feynman, Toffler and Webb. What we know is that we'll have narrow AI tools to dramatically improve the quality of human life for this generation while we advance its scope, and still we will ponder on the utopia, or fear of the potential general AI singularity that might follow.

Anand Srinivasan,
Global TMT Sector Leader & Senior Technology Analyst
at Bloomberg Intelligence, Pennsylvania, USA

Introduction

It takes something more than intelligence to act intelligently.

—Fyodor Dostoevsky (*Crime and Punishment*),
Russian Novelist and Philosopher

The first problem in trying to understand AI today is that we can't even agree on what it is. Open a dozen search results for 'What is AI?', and you'll find at least a dozen different answers. Most will probably be correct, but many will include technical language or scientific notation and be hard for a layperson to fully follow.

Many experts would argue that understanding AI and ML needs a strong understanding of mathematics, physics and computer science. If your aim is to build AI systems, they're absolutely right.

But you don't need an engineering degree to own or drive a car, or even work in the automotive industry. In the same way, non-specialists can understand and use AI. This book is for people who want to understand AI, how it works and what it means for us but don't have expertise in maths, science or technology.

WHY READ ABOUT AI AND ML?

AI and ML, and related terms such as 'algorithms' and 'deep learning', appear all around us. From daily newspapers to annual earnings calls, people use them freely and frequently, sometimes interchangeably, often superficially.

Typical discussions about AI with businesses include distinguishing between hype and reality and using AI to improve business results. Essentially, there are variants on 'how and when should we use

AI and ML in our business?' In contrast, conversations with friends and family on AI polarize into two big topics. The most passionate discussions are about concerns and threats, such as 'Big Brother' and privacy worries. But there's also frequent surprises at the latest conveniences added to daily life. For example, shopping coupons from stores that seem to know exactly what we need to buy this week, or novel features like self-parking in a new car. Curiously, these conversations rarely start off about AI, as people often don't even realize that AI or ML is involved.

AI has been a long time coming, but it's finally reached some kind of maturity and is quickly spreading. It will soon be, and arguably already is, all around us, touching almost everything we do and experience.

Reading about AI and ML will help you understand how they're changing the world around you. If you want to make sense of jargon thrown around by both experts and the uninformed, this book will help. Reading about AI and ML will also keep you on top of how they're affecting life at work. If you want to avoid being flummoxed by a new breed of colleagues, consultants and management initiatives, this book will help. And if you're still early in your career, reading about AI and ML will help you figure out how to prepare for and achieve your professional goals, and where these technologies might fit on your path.

WHAT THIS BOOK DOES (AND DOESN'T DO)

This book provides you with the clarity on terms such as AI, ML and the like in practice for everyday life at work and home. This includes the implications and changes this technology brings. Chapter 1 starts the discussion by exploring the meaning of intelligence generally, not just artificial. It uses this to introduce the concept of AI and other important terms in the rest of the book. Chapter 2 provides context that's often overlooked today, describing how AI came about and why it had so many false dawns. It also explains why this generation of AI is different from those that came before.

To help cope with the natural instinct of experts (in all fields) to use jargon to protect against outsiders, Chapter 3 explains in layperson terms how AI and ML work. This includes the maths and science and what people do when building AI. These and other terms are explained in non-technical language. Chapter 4 illustrates how all this comes together in the home, office, shop and pretty much everywhere. This widespread use creates dilemmas for those in the charge of the countries, continents and corporations on which we are utterly dependent. Chapter 5 outlines some of these challenges and what they might mean in practice. Finally, Chapter 6 covers factors likely to shape what AI and ML might become in the future. This includes looking at what some well-known names in field of AI have to say.

This book is grounded in fact rather than opinion. It illustrates and explains AI and ML using generally accepted knowledge and the author's personal experience. It refrains as far as possible from offering views other experts would take issue with. The goal is to retain a measured, neutral voice on a topic that doesn't always lend itself to that.

The book doesn't tell you how to build AI or ML systems, or advise on what to do about it in your business. Nor does it dig into details that require deep technical or specialist knowledge.

It's a concise introduction to a complex subject. So, by design, it doesn't explore in depth the intricacies and nuances of the topics. Nor does it elaborate in detail on the 'whys' and 'hows' behind what it describes. If you're interested in further detail, there's a comprehensive set of references. They're a mix of academic, general and specialist references to more in-depth coverage of the ideas and information covered.

1 WHAT ARTIFICIAL INTELLIGENCE IS, ISN'T AND MIGHT BECOME

Sometimes it seems as though each new step towards AI, rather than producing something which everyone agrees is real intelligence, merely reveals what real intelligence is not.

—Douglas Hofstadter,
Pulitzer Prize Winning Author and Cognitive Scientist

With 'artificial intelligence' (AI) now being a part of everyday parlance, what we think the phrase means is more relevant today than ever. To understand AI, we should understand what intelligence means generally, and what distinguishes man-made from natural intelligence.

As we learn about AI, we'll come across many forms. We'll see that understanding them becomes easier if we're clear on the different types of intelligence we're talking about. Perhaps AI has reached its own 'Eskimo Words for Snow' moment. Claims that Eskimos have many more words for 'snow' than others have been mostly misreported. But the central concept remains, that when something is important to a group of people, the language describing it evolves, becoming richer and more nuanced. To understand AI, everyday terms may no longer be precise enough, especially to appreciate important subtleties or distinctions. Our first port of call is the familiar and deceptively simple-sounding word 'intelligence'.

WHAT DOES ‘INTELLIGENCE’ MEAN, HUMAN AND OTHERWISE?

Before getting into intelligence in computers, let’s first think about what we mean by the word ‘intelligence’ in its own right, how we use the term and what we mean by it. Psychologists and researchers have been arguing about this for over a century. We don’t need to resolve that debate, but we do need clarity on what we’re talking about.

How Do *You* Decide How Intelligent a Person Is?

When you meet someone, you generally form an opinion pretty quickly on how intelligent they are, especially relative to yourself. That may change over time, but your view of their intelligence is based on their behaviour. When you do this, and judge that a person is ‘intelligent’, what’s behind that view? What causes you to consider them more or less intelligent than someone else?

Perhaps it’s how much they know, or how much more they perceive than you do. It might be how articulate or insightful they sound. Or maybe they’re just more successful at getting things done than others. There’s not a single, simple answer to explaining what we mean by intelligence in people. This one word covers a variety of characteristics.

A cursory attempt might uncover features like: (a) has more knowledge than us; (b) has more experience than us; (c) has a good memory; (d) makes insightful observations; (e) displays sound judgement (or perhaps wisdom); (f) thinks quickly; (g) solves problem easily; (h) achieves goals consistently; (i) learns from past experience; (j) adapts to new situations; and (k) creates new ideas and concepts.

In daily life, we can’t necessarily assess someone’s intelligence from a score or measurement. It’s more likely we’ll make a judgement based on how they display characteristics like those above. If there were a contradiction between our judgement and such a score (e.g. IQ or exam grades), we might well put more faith in our own opinion.

In other words, in normal usage, ‘intelligence’ isn’t an absolute term that can be defined (usefully) in a single way. It’s a highly contextual term that can vary widely in meaning. When you use it about a person, it’s usually a subjective reflection of what intelligence means to you personally, not according to textbooks or test results.

What About Intelligence in Other Species?

To discuss intelligence in machines, let’s first look at intelligence in other non-human creatures. You’d probably accept that humans are more intelligent than parrots and chimps. And definitely more than pigeons, dolphins and ants. Of course, that’s obvious, but nevertheless it’s useful, for our purposes, to try to articulate why it’s obvious.

Humans know more about subjects than animals and can process more complex ideas. We deal with abstract concepts that the other species can’t comprehend. With a few exceptions, we’re the only ones to farm food, play for fun and create art. Another difference demonstrating our superior intelligence is our widespread use of tools. Not just tools for efficiency, but also to enhance natural abilities and overcome limitations. For example, paper and pen to improve memory, levers to increase strength, and weapons to defend against predators.

But, before getting carried away with our clear superiority, we need to remember that for certain activities we’re inferior. For example: pigeons find their way back home over hundreds of kilometres, without maps, landmarks or GPS; dolphins use sonar to swim and hunt in dark, murky conditions, without light or night vision aids; ants build city structures with ventilation, sewage and transport networks, without machinery, plans or engineering degrees.

Does that make us less intelligent than pigeons, dolphins and ants? Clearly not. But it does mean that being more intelligent doesn't mean we can do everything better. More to the point, something less intelligent than us may be able to do some things better than us, things we may find difficult or even impossible, even with tools.

Now that we have a perspective on 'real' intelligence, we can move on to how it's used by those creating artificial versions. We start with the question of how to judge the presence and degree of intelligence in something.

Breaking Down the Concept of 'Intelligence'

When we looked at how a normal person judges intelligence, some characteristics were listed to consider, based on common sense. Researchers have built this concept, incorporating the work of psychologists, neuroscientists and other experts on the human brain. There's no universal, definitive list of what all scientists will agree to mean intelligence, but there are a few characteristics common to the major theories and frameworks.

Several researchers have adopted these for AI work, and the consensus of these experts[1] is that 'intelligence', artificial or otherwise, means the ability to display at least one of the capabilities listed in Table 1.1, often several.

Table 1.1 Nine Characteristics of Intelligence

Characteristic	Description
Reasoning (problem-solving)	Consciously make sense of surroundings, apply logic and adapt actions based on new or existing information
Perception	Be aware of, interpret and extract meaning from sensory information, both presented directly and in the surrounding environment

Characteristic	Description
Natural language communication	Communicate using language that has evolved through use, as opposed to artificial or constructed language
Motion and manipulation	Ability for something to move itself or move and control objects
Learning	Gain knowledge or skill by study or experience, including to improve the performance of an activity
Representing knowledge	Depict information about an item, activity or surrounding environment and ascribe meaning to it
Planning	Create strategies or sequences of actions to achieve an intended goal
Social awareness and skills	Understanding the reactions or likely reactions of others when interacting with them, and modifying behaviour accordingly
General intelligence	The integration of intelligence capabilities to solve new, unexpected or undefined problems

This list helps us understand AI, because it gives us characteristics to look for in a machine, computer program or other artificially created system, to decide if it's intelligent. The more these traits are present, the more intelligence there is. Conversely, and as importantly, if none of these appear, it's not intelligent, even if it initially appeared to be. The same applies to natural intelligence. For example, applying this to the earlier animal examples, parrots, chimps, pigeons, dolphins and ants, all exhibit several of these characteristics, and so can be classed as intelligent.

We now have clarity on a more useful meaning of intelligence, and a way to identify its presence. This gives us a way of deciding, relatively objectively, if a piece of technology can be considered intelligent and is therefore AI.

ILLUSTRATION: HOW INTELLIGENT IS A SMART SPEAKER?

We now have our first tool in our arsenal for understanding AI. We use it when presented with something that is claimed to contain AI. Its value comes from considering which of the characteristics of intelligence it displays, and how core they are to its purpose. Let's illustrate this with 'smart speakers' like the Amazon Echo.

These have become mainstream in the home and are basically speech-enabled computers built into music speakers, connected to the Internet via Wi-Fi. They play music, control devices connected to the Wi-Fi (like heating, lights and home alarm systems), and provide information they have access to, such as Internet, online calendars or data from connected devices. They consist of an electronic device (usually a music speaker with built-in microphones), equipped with software that can communicate using speech (Amazon Alexa, Google Assistant, Apple Siri).

Without doubt, they are intelligent devices. What's not necessarily obvious is that controlling devices around the home isn't the intelligent part. The home automation aspect is simply turning on lights, setting burglar alarms, turning down radiators and so on. These are simple electronic signals sent to household devices in response to an instruction, and none of the intelligence characteristics we listed are needed or used.

Their main piece of AI is an outstanding natural language capability. These devices are at the leading edge of the mainstream technology to hear and understand normal human language. We can't yet hold wide-ranging conversations with smart speakers, but we can talk to them in a human-like way about the things they can do.

Behind the scenes, these speakers also use AI that resides in things they connect to. For example, if you ask what the weather will be like tomorrow, it's not the smart speaker that is using intelligence that finds the answer, it's intelligence in a weather forecasting application (located on the Internet). That application uses sophisticated AI to analyse and predict weather and can pass results to the smart speaker. But the speaker doesn't possess any weather forecasting intelligence, even though it seems to know what tomorrow's weather will be.

The important point is that the smart speakers don't use their own intelligence to find the answers to such questions. They pass on requests that they neither understand or know how to answer (your instructions or questions), and speak back any result they're told to provide. They treat weather reports, music tracks and even jokes as pieces of data to pass to you. They don't 'know' what they're telling you, just that it matches a request you made for a piece of information. Their intelligence is in translating what you said in plain language into instructions other computers can respond to.

So, home automation devices like these are great examples of AI and demonstrate the Natural Language Communication intelligence characteristic very well. But beyond that they're actually pretty 'dumb'. Amazon, Google and Apple might counter that these devices are part of a wider set of devices and applications that work together to provide a lot of intelligence, known in the technology industry as an 'ecosystem'. In this case, the ecosystem of devices and applications working together to create an intelligent environment is called a 'smart home'.

This is a valid and important point and applies to much AI around us. When we use AI, much of the intelligence is distributed around a set of devices, services and applications. We may not be aware of that, because, as with the smart speaker, we may only be engaging with one of the pieces of that wider ecosystem.

Being Intelligent versus Being Good at Doing Something

The next step in understanding intelligence and AI is to look at the second earlier idea: the relationship between intelligence and performing an activity. We've previously established that how well something is done isn't an indicator of intelligence. We'll now explore that further, looking at the use of different types and degrees of intelligence to perform the same activity. We'll start with the activity of moving towards a desired destination. An extreme example is a plant 'deciding' to move towards the Sun. It does this through the biological process of phototropism, obviously using no intelligence.

In contrast, a pigeon decides which direction to move through mechanisms not fully understood, but which are obviously intelligent. The degree and type of intelligence depends on *how* it makes the decision on which way to fly. The three leading theories on that decision involve sun detection, magnetism and infrasound. Each involves differences in how a pigeon's brain uses its knowledge of where it is and where home is to decide which way to fly. All require Perception and Motion, but it's not clear if Planning or Reasoning are involved, especially if it's a purely 'instinctive' skill.

The third example, a person, achieves the same outcome using a very different form of intelligence. The closest example to the pigeon is walking home from somewhere nearby, say a neighbour's house. For such a familiar journey, the activity is almost automatic, and doesn't need much intelligence. But for a journey needing transport, or on an unfamiliar route, other intelligence characteristics kick in. Map-reading, programming a Sat-Nav or planning public transport, each require different types and amounts of intelligence, all substantially more than the stroll from next door.

So, the same generic activity performed by three species requires three very different forms of intelligence. In the human case, the same activity can use different types of intelligence, depending on our choice of solution.

In other words, we can't assess the intelligence used to perform an activity by simply looking at the result. We need to know more about how the activity is performed. What's required is a more granular appreciation of how the activity is broken down into different tasks, only one or a few of which may actually require the intelligence.

Illustration: How Intelligent Is a Simple Chatbot?

If you're unfamiliar with chatbots, they're the text windows that open up on many modern websites, inviting you to ask questions. Originally, these were a way for customers to communicate with customer service staff by typing instead of calling on telephone. Today, there's rarely a person at the other end, only a computer. It uses AI to answer customer queries by understanding what they want to know, and presenting relevant answers, such as about products or orders.

Basic chatbots are pretty simple pieces of technology. They're designed around a fixed set of answers (e.g. FAQs, product descriptions, etc.). For each option, there's a series of words to look for in questions. For example, if a question includes 'payment options', the chatbot might reply with a pre-written answer that lists different ways to pay. For groups of answers, the chatbot is designed to reply with further questions to clarify. For example, if a question includes 'skirt', the chatbot might be programmed to ask 'what colour skirts are you interested in?' or 'would you like to look at long, short or mid-length skirts?' The chatbot also requires logic for questions it can't handle. For example, if a product isn't sold or the customer asks what the weather is like, it needs appropriate answers, expressed as rules.

The 'rules' that tell the chatbot what information to provide or what clarifying questions to ask, are standard computer programming.

There's no intelligence, just a set of logic statements in simple computer code. The AI extracts relevant information from the natural language statements typed by the customer and decides which rules to apply. It's not as sophisticated as the intelligence used by smart speakers, because the activity is much more limited in scope.

A simple chatbot isn't particularly intelligent, because it doesn't need to be. If it's well designed, it will be ready for most questions people are likely to ask and will rarely reply that it doesn't understand. It's only if people ask about things it's not designed to handle that the limitations become apparent. A common experience of chatbots is that anything beyond a narrow set of questions generates an unhelpful response such as 'Sorry, I didn't understand that'.

That's all we're going to cover in the chatbot example here, although we'll return to it later. It's now time to return to our discussion of what AI means.

DIFFERENT TYPES OF AI

The examples of AI we've seen so far are representative of all AI today: They do a relatively narrow set of things intelligently, but don't work well (or at all) for anything beyond the activity it was designed to do. This is what AI experts mean by 'narrow' intelligence.[2] It's the opposite of human intelligence, which by contrast is called 'General' intelligence. These are sometimes also referred to as 'weak' and 'strong'. We'll now look at what narrow AI is already capable of.

What 'Narrow' AI Can Realistically Do Today

The amount of intelligence required to perform narrow AI varies significantly, even for the same type of activity. For example, in the two natural language communication examples, the smart speaker uses much more intelligence than the Chatbot for communication. In contrast, general AI is an AI that can be applied to a wide variety of activities. We'll see later that human intelligence is the benchmark that researchers aspire to reach with general AI.

To understand what's going on in a given piece of AI, we need to break down the activity into components, and identify the intelligent pieces involved. Those pieces are, in effect, the building blocks of intelligence, used to construct more complex AI systems and solutions. Today, and for the foreseeable future, even the most complex AI system will remain narrow. No matter how sophisticated the AI used is, any examples we see today will only be able to perform one activity (which may include a set of activities), and won't be able to cope with anything else it wasn't designed to do.

Table 1.2 contains some common AI building blocks used today to create artificial narrow intelligence (ANI) applications and systems, which may contain several such components. This list illustrates the kind of individual activities AI can do, and so helps you spot, or at least suspect, how AI is being used.

Table 1.2 Narrow AI Building Blocks

Intelligent analytics	Intelligent search
	Forecasting and prediction
	Anomaly detection
Computer vision	Image processing and recognition
	Text and handwriting recognition
	Video processing
Natural language processing	Natural language generation
	Natural language recognition
	Sentiment analysis
Intelligent automation	Virtual assistance/RPA (Robotic Process Automation)
	IoT (Internet of Things)
	Robotics
	Autonomous vehicles

Now that we've seen what ANI is, we will look at general AI, which will also give us a sense of how AI can be compared with human intelligence.

'General Intelligence': The Ninth Characteristic

General intelligence was the ninth characteristic of intelligence listed in Table 1.1. It describes the kind of intelligence we take for granted in humans. It's the ability to face a new situation and figure out how to deal with it. The main, and as yet insurmountable, challenge in creating an artificial general intelligence (AGI) is the infinite variations in circumstances it might face. When humans deal with this, it's not a conscious exercise, and scientists are yet to really understand it.

For example, if a stranger approaches and asks a question, we make an instant decision on whether to answer, ignore them, hold our bag more tightly or many other possibilities. We don't yet know what our brain does during the instant we're figuring out what we're dealing with and deciding how to respond.

Because we don't understand it; we can't break down the first part of a general intelligence activity into a set of smaller tasks. Hence, unlike ANI, we can't yet design an artificial way of emulating it. Defining it is also elusive,[3] and without definitions, scientists struggle to frame useful questions and build models that lead to solutions.

Some experts believe that understanding, modelling and recreating general intelligence in AI will never be possible. Others disagree, and it's an interesting debate. The fundamental difference of opinion is whether it's inherently possible to understand and model the human brain, and therefore human general intelligence, well enough to recreate it artificially.

AGI would require a huge step forward in our understanding of how the brain works. We'd then need to build an artificial version. But if that could be done, an artificial brain should, at least theoretically, exhibit the same characteristics as a human brain, including general intelligence.

There are many reasons to consider this unlikely to ever happen, such as around computing power and the biochemical processes. But if those overcome, then general intelligence could theoretically be possible. So, the argument goes,[4] because general intelligence could theoretically be achieved, it is in principle possible, regardless of likelihood or timescale.

The counterargument is based on comprehensive observations about the complexity of the human brain, our attempts to understand and model it, and the difficulties encountered so far. The other side of the argument is built on the assertion that these observations make it clear that the challenge is overwhelmingly difficult, and for all practical purposes unachievable.[5]

Both viewpoints are couched in hypotheticals and assumptions, and all we can really be sure of about general AI is that it's incredibly difficult, and if we do achieve it, it won't be anytime soon.

Simulating versus Replicating Human Intelligence

There's a fascinating variation of this argument from a subset of AI work called the Philosophy of AI. The subject remains whether general intelligence could ever be achieved by AI. But the debate shifts to whether AGI is something that simply needs to be displayed, or if it has to be achieved in the same way the human brain does it. This debate raises the question of whether there's something inherently different about the human brain compared to a computer, some special quality we might call 'consciousness' or the 'mind'.

A researcher called John Searle put forward two hypotheses[6] to distinguish between these two ideas. (Perhaps confusingly, he called them the Strong and Weak forms of the AI hypothesis). Searle became a contentious figure, but his statements below very neatly summarize the underlying point.

AI hypothesis, strong form:
An AI system can think and have a mind.

AI hypothesis, weak form:
An AI system can only act like it thinks and has a mind.

Those who support the first statement argue that general AI will require computers to possess a mind, or consciousness. Proponents of the second believe that general AI will be achieved when computers behave as if they have general intelligence, and it's irrelevant whether this is done exactly the same way as the human brain.

The importance of this distinction is its relevance to the long-term goals of AI research. It determines whether AI is about *simulating* human intelligence or *replicating* it. If it's the former, then what matters is that AI achieves results as good (or better) than a human doing the same thing. For the latter, we need to make AI work in the same way as the human brain.

Outperforming Human Intelligence

So far, we've identified two types of AI: 'narrow' and 'general'. AI scientists also include a third type: artificial superintelligence (ASI). This is AI that exceeds human intelligence, and is what AI development will strive for after AGI is achieved.

Even though ASI is furthest away from reality, it's perhaps the type of AI that comes to mind first for the general public. That's because authors and movie makers, unconstrained by the limitations of today's science, have brought to life many fictional examples. The humans in *Star Trek*, *Terminator*, *Star Wars*, and *2001: A Space*

Odyssey, live and work alongside computer and robots with far greater intelligence than they have.

We won't explore ASI further here, as it's too far removed from our goal of understanding AI. We'll come back to the wildly unachievable concepts of ASI later. For now, we'll stay with the merely mildly unachievable idea of AGI.

How Would We Recognize AGI?

For any specialist in a field, ambiguous, subjective language isn't helpful, especially for key terms. Scientists take things a step further and like to be able to measure important things. To claim progress, they need quantifiable, reproducible results that can be validated by others. So if an AI scientist believes they've created a machine that achieves general intelligence, they'd need to do two things: demonstrate clearly why their claim is valid, with unambiguous, verifiable evidence; and then repeat the exercise (or ideally allow others to) with the same result, usually as an experiment or demonstration.

AI researchers have approached AGI in a way that's widely accepted in other disciplines: setting theoretical 'tests' beforehand that create a bar that a successful piece of work would need to cross. They're defining up-front what a successful result will need to demonstrate. As the field has progressed, different aspects of general intelligence have become important. So, there are now several tests for AGI, all waiting for the day we have something to apply them to.

Below are four well known tests for general AI. They bring to life the contrast between narrow and general AI and provide insight into the minds of the people shaping the AI we'll be living with.

The Turing Test

Alan Turing[7] was a mathematician and computer scientist in the mid-20th century. He is regarded by many as the father of theoretical computer science, and by some as the father of AI. With the crude computing technology of the time, he could put little of his AI work into practice. That limited some aspects, but liberated others from practical constraints.

He made a significant contribution to AI with a hypothetical test for whether a machine's intelligence is on par with humans. The test was specific, but 'Turing Test' is often used as a generic term for tests of how close AI is to human intelligence.

The Turing Test involves a person conversing with an intelligent machine using keyboard and screen (so appearance or voice don't affect the test). If the person can't tell if they're communicating with a human or computer, the AI is considered general AI.

His specific test is a somewhat contrived version. It was presented to the public in a 1950 paper[8] that opened with the words 'I propose to consider the question, "Can machines think?"' That question suffered from familiar problems of ambiguous definitions (in particular 'think'). So, Turing went on to frame the discussion around a different, related question which better lent itself to scientific investigation. He did this by devising a theoretical game called the 'Imitation Game' and asked a new, related question: 'Are there imaginable digital computers which would do well in the imitation game?'

In the Turing Test, the imitation game is a three-person game played by human questioner, human opponent and computer, each in separate rooms. The questioner asks questions to both, and the answers are returned in writing. The goal of the questioner is to consistently distinguish between the person and computer from the answers. The computer's goal is to consistently persuade the questioner that it's human.

Perhaps ironically, the Turing Test has generated debate because of uncertainty around what he was trying to achieve in his paper.

The ambiguity is whether the Turing Test is about how well a computer could fool a questioner into believing it's human, or how well it can imitate a human trying to do that. In other words, did Turing intend to test a computer's ability to *win the game*, or its ability to *imitate a human* playing the game? The latter is generally accepted, but there isn't any universal agreement on the answer.

The Coffee Test

The first well-known alternative to the Turing Test was the Coffee Test. In 2007, general AI was covered in a PC World magazine[9] interview with Steve Wozniak, Apple co-founder. In typical Apple style, Wozniak presented something that had previously been difficult to understand, cutting through complexities and stripping out jargon.

He suggested that a real example of general AI should be able to enter a typical American home and make a cup of coffee. Deceptively simple for most humans but requiring advanced levels of all nine intelligence characteristics.

As with the Turing Test, the Coffee Test is built around a single activity, but with complexities that make it too difficult for narrow intelligence. Here, the single activity is making a hot drink rather than playing a party game. What makes it more than narrow intelligence is not making coffee but figuring out a variety of other things along the way.

For example, simply entering a home and getting to the kitchen isn't straightforward for AI (which we assume can move around freely). Once there, it's no big deal for a person to find coffee, mugs, spoons and so on. But for AI to do that, it's a very big deal.

The Robot College Student Test

In 2012, AI researcher Ben Goertzel published an article in *New Scientist*[10] about conscious thinking in machines. He proposed a new type of general AI test, as he thought the Turing Test focused too narrowly on imitating human conversation. He felt that a better test should be for a conscious robot to achieve a substantial goal that humans routinely manage.

He suggested the Robot College Student Test: 'When a robot can enrol in a human university and take classes in the same way as humans, and get its degree, then I'll consider we've created a human-level AGI: a conscious robot'.

There isn't much more to say, as the reason why this is a much tougher test of general AI is self-evident. But as with the Coffee Test, thinking through what a human does to pass the same test, and considering how we'd begin to break it down into smaller steps and automate them, will hopefully reinforce why general AI is such a challenge.

Perhaps surprisingly, there wasn't much response from the research community. It's nothing like as well-known as the Turing Test, which is disappointing, as it's an intriguing idea. It did spawn some humorous discussions, such as whether such an AI would realize the inherent futility of getting a degree and question the authority of the humans who built it!

The Employment Test

Our final example of a test for general AI was proposed by Stanford Professor Nils John Nilsson in *AI Magazine* in 2005.[11] It's not a pass/fail exercise, but a spectrum of capability. It doesn't just show when general AI has been reached, but progress towards it.

Nilsson's test is built around his assertion that: 'Machines exhibiting true human-level intelligence should be able to do many of the things humans are able to do. Among these activities are the tasks or "jobs" at which people are employed'.

His test is whether AI can perform the jobs ordinarily performed by people. These of course vary tremendously, and using it in practice would need conditions, caveats and clarification. Nilsson addressed this by picking specific jobs from 'America's Job Bank'. These included Maid, Convention Planner, Tour Guide and Security Guard.

Nilsson went on to explain that 'Progress toward human-level AI could then be measured by the fraction of these jobs that can be acceptably performed by machines'.

As general AI is unlikely to be achieved in the near future, we won't spend more time on it. But hopefully, exploring how we would recognize general AI if we saw it adds clarity to our understanding of the complex meaning of the term AI, especially if it's described as Narrow, General or Super.

A NON-TECHNICAL GLOSSARY FOR AI AND MACHINE LEARNING

Understanding AI and ML requires familiarity with new terms, and this section explains what the most important of them means. Many of course sound technical, but several have been annexed from regular language.

The Basics of AI and Machine Learning

Artificial intelligence (AI)	A computer-based system that simulates characteristics of the human brain in order to perform activities that could otherwise only be performed by humans.
Machine learning (ML)	A subset of AI with the ability to gain new knowledge from experience and perform activities not explicitly defined in its design or programmed instructions.
Explainable AI/XAI	AI which shows how a result was calculated, so that the basis for a result can be examined and explained. Normal ML doesn't automatically do this.

Big data	The analysis, processing and use of large volumes of data that are too great for traditional data analysis tools like spreadsheets and regular data management tools.
Deep learning	A type of machine learning that uses several steps, or layers, to calculate a result.
(Artificial) neural network	A set of ML deep learning layers that perform all the steps required for an entire ML calculation. Each layer contains several parts, each of which may connect to several others, hence the name network. Called neural because it loosely mimics how the human brain works, using artificial versions of a biological entity called a neuron.
Supervised learning	A type of ML trained using data that contains labelled examples of the desired results it should generate, so that it can improve the way it works until it can obtain the same results itself.
Unsupervised learning	A type of ML trained using data that doesn't identify the desired results in a set of data either because the desired results aren't known or suitably labelled data isn't available.
Reinforcement learning/Semi-supervised learning	A type of ML which solves a set of individual problems so that they collectively achieve a larger goal, for example making individual moves in a board game that lead to winning the overall game.
Transfer learning	A type of ML which applies the learning gained by solving one problem to a different one.
Accuracy/ Sensitivity/ Precision/Recall/ F-Score	Different terms to indicate how well AI obtains results, akin to the probability of a right answer. For example, the likelihood of a patient having cancer based on a test result. Each term represents a different type

	of correctness appropriate for different circumstances. For example seeking all results that might be positive vs only wanting those that are definitely positive. Related to terms such as false positive and false negative.
Algorithm	The mathematical calculations performed on data to generate a result from AI.
Vector/Matrix/ Linear algebra	Vectors and matrices are tables of data, and linear algebra is the maths used to perform calculations on them. They allow large sets of numbers to be handled quickly and efficiently, and are used in AI to handle calculations on large volumes of data.

Algorithms and Models

Dimensionality reduction/ Classification	Statistical technique to find structure in sets of data by finding groups, categories and relationships. Used to simplify and organize large amounts of data before performing further analysis.
K-means clustering/ K-nearest neighbours	Two common techniques to find groups of similar data in a set. Clustering is used in unsupervised learning to find naturally occurring groups. Nearest Neighbours is used in supervised learning to quantify the likelihood that new data points are correct.
Gradient boosting/Gradient descent/Stochastic gradient descent	A set of ML techniques to improve ML models by combining several to minimizes the errors of each. 'Gradient descent' is the use of a 'loss function' to reduce errors; 'Stochastic' means randomness is used to improve accuracy; 'Boosting' combines models to improve results.

Regression/ Linear regression/ Logistical regression	Regression is a statistical technique commonly used in AI to predict a value from various factors. There are many forms, the two most common being logistic (two possible outcomes) and linear (many outcomes).
Support vector machine	A widely used ML algorithm in supervised learning for classification and regression tasks. Similar to clustering techniques like K-means.
Decision tree/ Random forest	Another supervised learning algorithm for classification and regression. Creates 'rules' during training to create a mathematical version of a flowchart. Random forest is the combination of several decision trees to create a more complex algorithm with better performance.
Backpropagation	The passing 'back' of information through layers in an ML neural network to allow feedback about events that have previously happened to be used for future calculations.
Deep belief network/Deep autoencoder	A deep belief Network is a type of ML algorithm consisting of several layers, often used in computer vision systems. A deep autoencoder is similar, and comprises two deep belief networks.
Naïve Bayes classifier/ Bayesian networks	Rev. Thomas Bayes was an 18th-century statistician who developed a theorem to model how humans use evidence to convert beliefs into predictions. Naïve Bayes classifiers are ML algorithms that use his theorem to classify data. Bayesian networks use it to model the relationship between cause and effect in data.

FFN (feedforward network)/Multi-layer perceptron	FFNs are deep learning neural networks that pass the results forward during calculations. The multi-layer perceptron is one of the oldest FFNs, but still widely used.
Attention mechanism/CNN (convolutional neural network)/ RNN (recurrent neural network)/ LSTM (long short-term memory)	Many neural networks require feedback between layers, as calculations in one layer depend on previous results. Attention mechanisms are ways to store information during calculations for reuse by other layers. CNN and RNN and LSTM are neural networks that do this.
GAN (general adversarial network)	GANs are best known from their use in creating 'deepfakes'. They are ML algorithms consisting of two neural networks that work against each other to produce a result. One creates new data based on an existing pattern (e.g. a photo), the other tests it for authenticity and rejects it if not good enough.

2 A LONG AND TROUBLED LIFE STORY

The error of youth is to believe that intelligence is a substitute for experience, while the error of age is to believe experience is a substitute for intelligence.

—Lyman Bryson, American Educator and Author

To understand AI today, you don't need to know its life story, but it's a useful and interesting piece of knowledge. An appreciation of its origins helps understand some of the concepts underpinning modern AI, many of which date back many centuries. AI is a technical subject today, but its roots lie in metaphysics, and much of its history was written by philosophers, storytellers and even religious leaders. In more recent times, it's been the domain of physicians, psychologists and science-fiction writers. It was only in the late-20th century that it became a science.

The story began with early civilizations, grappling with big questions about who we are, where we came from and why we exist. Religion and folklore suggested some answers, but in doing so also posed other questions. Many remain unanswered and continue to challenge us. Let's now take a look at this life story, at how AI and ML were conceived, born and raised. Like all good stories, we'll start at the beginning.

A GLEAM IN THE EYE OF EARLY HUMANS

Ancient Civilizations

The major ancient civilizations included some version of the concept of AI in their mythologies, religions or philosophies. The roots often lay in attempts to explain how the human race came to be, which invariably meant the existence of a higher form of intelligence. If you accept the principle of a greater intelligence than humans, you also accept that intelligence isn't restricted to humans. And if human intelligence was created by a higher being, a natural question is whether humans could in turn also create intelligent beings.

Ancient Greeks were believed to be the first to attempt to create intelligent beings—sacred statues claimed to possess minds, emotions and wisdom. However, the Greeks usually attributed creation of robots to the gods. The earliest on record were made of gold by Hephaestus, the God of metalworking. He went on to create a female robot called Pandora, better known for opening the box that released evils into the world.

Greek mythology is full of other references to what we'd now call robots. According to some translations, Homer's *Iliad* mentions robot servants. Daedalus, whose son Icarus died by flying too close to the Sun, was also credited with making moving, speaking robots.

The Greeks also recognized the threat of robots, and their potential as weapons. The best-known appears in the saga of Jason and the Argonauts, as Talos, the bronze warrior guarding the island of Crete. Talos, incidentally, was another creation of Hephaestus.[12]

Other robots appear in Indian, Egyptian, Chinese and Japanese writings. For example, Sanskrit writings describe how, after Buddha's death, his precious relics were protected by King Ajatashatru[13] near modern-day Patna. Rather than using traditional statues of warriors to guard them, he was said to have used 'spirit movement machines'.

The Middle Ages to Victorian Times

Such stories persisted through subsequent centuries across cultures. A central theme was a magical ingredient to imbue life into inanimate matter. Amongst the earliest writers on this subject was Jābir ibn Hayyān,[14] sometimes known as the father of chemistry. He wrote about alchemy (transmutation of matter), including Takwin: artificial creation of life.

Less scientific approaches to bringing statues to life were widespread. A famous one was the use of Hebrew words to bring life to a mud statue known as a Golem. These first appeared[15] in the Bible and Talmud, then in the Middle Ages in works of religious scholars.

By Victorian times, this had become the subject of fiction, the most famous being Mary Shelly's *Frankenstein*.[16] There were others, and they created alarm among readers. One person who put his concerns into writing was Samuel Butler.[17] In 1863, he questioned whether Charles Darwin's idea of evolution might apply to machines. He feared it would and predicted the replacement of humans by machines as the dominant species on Earth.

We close with Karel Čapek's 1921 Czech play *R.U.R.*,[18] set in a factory that makes artificial people, or 'raboti', a Slavic word that variously means slave, worker and drudgery. The title stood for *Rossumovi Univerzální Roboti*, translated into the English subtitle 'Rossum's Universal Robots'. Which is where the word robot originates.

CONCEIVED BY PHILOSOPHERS, BROUGHT TO LIFE BY MATHEMATICIANS, SCIENTISTS AND ENGINEERS

Mechanized Human Thought? Originally a Question for Philosophers, Not Engineers

As we've seen, from early times humans were fascinated by the possibility of creating the artificial life described in their myths.

This was originally the domain of philosophers, and their starting point was trying to represent thought—a theoretical, conceptual problem. With hindsight, we see that they were taking the first steps in solving an engineering problem: How to mechanize the activity of thinking. Those steps were initially taken by the great thinkers of Greek, Chinese and Indian societies. They developed structured approaches to performing, and so representing, logic, deduction and reasoning.

There were many well-known names involved, and some of the earliest realized that mathematics was a key part of this. Hence, Aristotle and Euclid were important. An equally important but less well-known name was al-Khwārizmī.[19] His major achievement was in the development of Algebra, but his name lives on as the origin of the word 'Algorithm'.

The shift from representing thought to designing a thinking machine happened in the Middle Ages. An early mention was by Spanish philosopher Ramon Llull.[20] He conceived of machines that could produce knowledge by logical means. He described them as mechanical entities, able to perform simple logical tasks mechanically. They performed these operations on 'basic and undeniable truths', what we might describe today as data. He didn't realize it at the time, but he was describing what would become the modern computer.

He had great hopes for what his machines could do and made immodest predictions for them. He claimed that such an approach could be used to produce all possible knowledge in the world. This might be a good time to remind ourselves of Google's mission[21]: to 'organize the world's information and make it universally accessible and useful'.

Philosophers Recognize Mathematics as the Language of Thought

Llull's work was influential, particularly for Gottfried Leibniz, who speculated that such mechanical calculation could be used for more than just creating knowledge.[22] He suggested it could be used to replicate reasoning as well. He worked on this in parallel with Thomas Hobbes and René Descartes, who were exploring whether all rational thought could be systematically represented. If it were, mechanizing it would be the natural progression.

One of their ideas was that algebra and geometry were the tools to do this, and they've been proven correct. In 1651, Hobbes distilled this concept into 10 words in[23] The *Leviathan*: 'reason … is nothing but reckoning (that is, Adding and Subtracting)'.

They conceived of a single language which could represent reasoning and logic as unambiguously as calculations. They believed this would turn philosophical disputes and debates into objective determinations. Leibniz described their ambition in concrete terms:[24]

'There would be no more need of disputation between two philosophers than between two accountants. For it would suffice to take their pencils in hand, down to their slates, and to say to each other….: "Let us calculate"'.

Mathematicians Describe AI Using Maths (Theoretically)

Between the Middle Ages and the early 20th century, the baton was handed from philosophers and alchemists to mathematicians and scientists. They were still pursuing the goal of mechanizing human thought, but the problem had fragmented and expanded.

The issues were not simple, and it would be another three centuries before the theories were brought to life. The first major milestone was George Boole's *Investigation of the Laws of Thought* in 1854.[25] In it, he expanded on two mathematical topics that are central to AI today: logic and probability. What makes it especially interesting for AI is that he was attempting to present mathematical proof for 'an intelligent self-existing being'.

Another important contribution was Gottlob Frege, who half a century later wrote a book called *Begriffsschrift*.[26] This was the most complete attempt of its time to create a full language based on mathematics to fulfil Leibniz' ambition of representing all logic and reason. Russell and Whitehead made substantial progress on this in their three-volume work *Principia Mathematica*.[27] They simplified and added precision to the earlier work of others, and addressed many of the paradoxes and criticisms that had not yet been resolved.

Finally, we'll close with the contributions of David Hilbert and Kurt Gödel in the 1920s and 1930s. Hilbert was known for his work on formalism, a branch of the philosophy of mathematics that added to the treatment of mathematics as a language. He argued the importance of interpretation and semantics, extending it to wider systems in which everything could be described.[28] We place this alongside Gödel because of the latter's Incompleteness Theorem.[29] That achieved many things, but unfortunately for Hilbert, it also showed limitations in his earlier work, in particular the use of mathematics as a language.

Together, these writings influenced subsequent AI work. They set limits on how mathematics could be used as a language to represent systems of thought and reasoning, then showed how, within those limits, it could be used to design mechanized thinking.

Computer Scientists Build the First Thinking Machines Philosophers Had Envisaged

We've now reached the Second World War, which brought two things for AI. First, it drove scientific research in a way never before seen. Countless achievements and breakthroughs came about between the 1930s and 1950s, much of it funded by governments. The second major feature was the rise of America as a world power, particularly in technology and science. The US Department of Defense started to wield influence on funding and prioritization of

US research, particularly through the Defense Advanced Research Projects Agency (DARPA).

British, French and German scientists of the day also continued to make great strides in developing computers as machines that could think. There are many histories of computing that describe the frenetic activities of the period,[30] from the code breakers at Bletchley Park in the UK to the creators of the ENIAC computer in the USA. These were built on the work of many others, including Charles Babbage, Ada Lovelace and Herman Hollerith.

The result was a new type of machine: the computer, a version of what the philosophers of early times had imagined. Scientists were finally seeing the first glimpses of what would one day become thinking machines.

Spurred on by this, some scientists started work on a field we now know as AI. They represented diverse disciplines and perspectives, and a loose community of like-minded individuals formed. The official birth of AI was the first formal meeting of that group.

The Birth of AI Is Registered in June 1956

By the mid-1950s, a Maths professor named John McCarthy[31] at Dartmouth College, New Hampshire had developed an interest in the emerging but yet unnamed field of AI. He had found kindred spirits across the USA and the UK, at places like MIT, Bell Labs and IBM. In 1955, he secured funding for a summer-long workshop[32] to study a set of problems in the area of thinking machines.

He developed the ideas with Marvin Minsky, Nathaniel Rochester and Claude Shannon. They are credited with introducing the term AI to describe a field of study that mostly didn't yet exist.

The group met in Dartmouth for around two months in the summer of 1956, and their goal was ambitious:

> The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.

One of the achievements of what became known as the Dartmouth Conference was agreeing to the term 'artificial intelligence' as the name of this new discipline. Over the summer, the conference defined the initial and long-term goals of the field, achieved its first successes and established its founding fathers. They recognized that if a longer term ambition was creating artificial brains, broad expertise was required. So, they included psychologists, economists, political scientists and engineers as people to work with.

The Dartmouth Conference marked the beginning of an exciting, productive and optimistic period in the early life of AI. Progress was made in every conceivable area, and hopes were high that results would follow. In many ways, the world of AI during the late 1950s and 1960s reflected wider changes in society, certainly in the USA and Europe.

And like the rest of the world around it, the swinging 60s were followed by the 70s, when things started to feel very different, as bleaker times started to make themselves felt.

KEPT ALIVE BY HOLLYWOOD THROUGH THE 'AI WINTER'

After the optimism of post-War AI progress, the 70s was a time of relative frustration and disappointment for AI. With hindsight, perhaps short-term disappointment was always unavoidable after such literally super-human ambitions.

Maybe it was also a sign of the changing times. There was progress on many fronts, including technology, without the shadow of global armed conflict. But memories faded, and a new form of global conflict started, the Cold War. This was played out by scientists as much as soldiers, with nuclear research and space travel among the battlegrounds.

Politicians became overtly interested in technology, bringing pressure for results that would translate into geopolitical or electoral advantage. There was also impatience, after years of unrealistically high expectations, matched only by the funding that went with them.

The questions from governments and funding agencies about AI got tougher, and funding became more difficult to come by. There were still breakthroughs during the fallow period that followed, and work continued on AI that we use today, including speech recognition, self-driving cars and computer vision. But despite some important seeds being planted during the next couple of decades, those years became known as the 'AI winter'.[33]

The AI Winter: Lucky Escape or Healthy Pruning?

Labelled as a single period, the AI winter was actually two: 1973 to 1980 and 1987 to 1993. Much effort was wasted, including earlier projects being stopped short of completion. Some new projects were stillborn, as unrealistic conditions were set; others were allowed to proceed, but with no realistic chance of success.

But in nature, winters are harsh times that precede periods of growth, and difficult weather conditions are an environment for clearing out that which is no longer needed. Much activity ceased in the AI winter, but the strongest restarted alongside fresh initiatives.

The Public Gets to Know AI: On TV and Cinema Screens

While AI researchers struggled to find funding, Hollywood maintained its fascination with the story of machines challenging human supremacy. The space race also kept imaginations excited, and much fictional AI existed in space. So, despite AI work not making as much progress as hoped for, it became more familiar to the general public.

Directors and screenwriters seemed torn between disaster and happy endings, with a mix of predatory machines and obedient electronic helpers brought to life. This had an effect on general attitudes and expectations towards AI. The concept became well-known, and terms like robot entered normal language. After watching so many fictional intelligent computers and humanoid robots, people also felt they understood what AI would look like when it became real. However, because the rest of the picture was usually so futuristic, they also automatically assumed that AI would only arrive far in the future.

SENT OUT INTO THE WORLD TO EARN A LIVING

By the time the AI winter ended, the world had changed beyond recognition, not just in terms of technology. By the 1990s, the fall of Communism had been accompanied by a near-universal acceptance of free markets and globalization. With this came the rise of private and institutional funding as a driver of technology development and innovation.

This had several implications for researchers in all fields, but particularly technology, and AI specifically. The most significant was the increased focus on financial returns from technology investment. Previous eras had been heavily influenced by government money, and therefore longer term and national goals. In contrast, commercial funding came from institutions with shareholder obligations. Their timescales were much shorter, typically needing to demonstrate annual achievements. Even longer-term investors, such as the emerging breed of venture capitalists were bound by 5–7 year horizons.

The impact of this on AI work was profound, with positive and negative effects. There was strong pressure for not just results, but results that would turn into revenue. And these revenues usually had to arrive in a 5-year horizon or thereabouts. Few organizations had the financial muscle to invest in technology that needed longer gestation.

The good thing about this was that researchers were forced to focus on the practical, commercial feasibility of their work, and AI technology was made to pay its way. The more challenging consequence was that it became harder to secure funding for some of the most important AI problems, especially those that would take much longer to reap commercial rewards. This meant that there was only a small set of companies and institutions with the appetite for longer term or speculative research. Many of these are household names today and have an advantage in their AI capabilities that few competitors can hope to overcome. The most well-known examples are Google, Facebook, Amazon and Microsoft. Others have joined them more recently, for example, Baidu and Alibaba. And of course, many leading academic scientific institutions are leaders in AI, such as MIT and Cambridge University.

This polarization started to level out as the funding market evolved. The same market forces that shaped the work being funded also influenced the investment available. As longer-term research led to longer-term commercial rewards, so the funding for such work increased. Today, academic and theoretical AI research is strong again, and AI research is better funded than at any time in history for both short- and long-term goals.

The result of all these changes on AI today is that how the technology pays for itself is now as important as how it works.

This brings an end to our recounting of the life story of AI so far. If it were a regular biography, this would be where the subjects have just started their working life, with a lifetime of potential ahead of them. Along the way, they'll be faced with decisions and crossroads, each of which will shape their ultimate path and destination. The rest of this book is about helping us understand AI better, and what some of those decisions and crossroads might be.

3 HOW ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING WORK

Any sufficiently advanced technology is indistinguishable from magic.

—Arthur C. Clarke, Science Fiction and Science Writer

The idea of understanding how a machine thinks and learns might seem daunting. If your goal is to build AI or ML systems, there's a lot of complex science to master first. But our analogy from the outset has been that you don't need to be an engineer to drive a car or even understand how cars work, and it holds true for AI and ML.

This chapter describes the most important principles, techniques and technologies involved in the creation of thinking machines. It builds on some of the ideas described in the first two chapters and adds a couple of new ones. By the end of it, you still won't be able to create AI, but you hopefully won't be mystified by how it achieves its results.

As with the earlier discussion of intelligence, some of the ideas are a little abstract, and in isolation may not seem particularly useful. To help, illustrations put the concepts into practice. You'll also need a little patience, because a holistic understanding of AI requires understanding several separate components and building blocks first.

To do this, you may find yourself returning to earlier pages. You might also see that something previously covered may take on a different meaning when seen alongside a new idea. Because of this, there's deliberate repetition of some points, as a reminder or elaboration.

SIX KEY CONCEPTS BEHIND AI AND ML

If you read a textbook or syllabus from any good ML course, even understanding the topics usually requires prior technical knowledge, including maths, statistics and computer science. This chapter covers many of the same topics without such foundations. To do this, we need an equivalent set of foundations for the layperson.

These take the form of six non-technical principles to underpin our understanding of AI and ML. These are not alternatives to maths or computer science, but a different perspective on how AI and ML work. They're designed to support our goal of understanding how AI works to demystify it, not build it.

These concepts are not something you'll find elsewhere and are not part of traditional AI education or practice. They are the result of my own AI research, and form a proprietary model I use in my AI work.

Concept 1: AI Needs Well-defined Problems, with Specific Boundaries

We saw earlier that AI today is narrow—the application of machine thinking to perform a single, defined intelligent activity. That activity may be very complex, and doing it may require several other pieces of AI working together, but it is still one activity.

A simple example is optical character recognition (OCR), a subset of a type of AI called computer vision. Printed text is digitally photographed, and an AI-enabled computer inspects the image to identify characters and convert them into digital text. The intelligence is in perceiving and recognizing the marks on the paper.

We also saw the gulf between artificial and human intelligence, how even simple things humans do can defeat AI today, such as making coffee in an unfamiliar house. Hence, general intelligence is currently impossible. Thus, all AI today is narrow, so AI today only applies to problems with clear definitions and boundaries.

Concept 2: Intelligence Means the Presence of At Least One of the Eight Characteristics

The second (hopefully) familiar idea in this chapter is the concept that intelligence is a complex term and is comprised of up to eight characteristics. These include reasoning, perception, natural language, motion, learning, representing knowledge, planning and social awareness. We are excluding general intelligence because we're only trying to understand how ANI works.

The converse is also true: The existence of at least one of the eight characteristics means that something is intelligent and exhibits Narrow Intelligence. This means if we want to understand how AI works, we need to understand how human-created machines can simulate these eight different intelligent characteristics.

Concept 3: Intelligent Activities Are Usually Comprised of Several Smaller Ones, Only Some of Which May Be Intelligent

The third concept is based on the idea that any complex problem can be solved by breaking it down into simpler ones. This applies to activities as well, so a complex activity is invariably composed of a series of simpler tasks. If the overall activity is an intelligent one, only one of the sub-tasks it comprises needs to be intelligent.

This principle helps you spot and clarify where the actual intelligence lies in an activity that AI supports. This can be a surprisingly formidable challenge, because we often do complex things without conscious thought, or even awareness, of the tasks it comprises. A simple thing like throwing a ball involves an incredible set of movements, calculations, decisions and adjustments. Catching is even more complicated. Yet if you were to ask people what they do during these simple actions, most would probably struggle to fully explain every aspect.

Concept 4: Data Is the Fuel of AI and ML

Our fourth principle isn't necessarily obvious from a comparison of human and AI. It's that AI and ML rely on data about the activity they perform. They need it to decide what to do, how well they're doing it, and how to do it better. Without this, what may initially look like AI may simply be automation. This principle is crucial to how AI works, and why dealing with data can be the most time-consuming part of building AI.

The reason it isn't necessarily obvious is because the huge amounts of data involved may have been used in the creation of the AI, but not necessarily its operation. For example, the advanced speech recognition in a smart speaker doesn't need to spend hours listening to a user to learn what they're saying, because it's already been trained. That training, done during development of the product, well before it reaches you, involved many hours listening to speech by a variety of voices and accents, so the smart speaker can already understand most customers from the outset.

Concept 5: Intelligent Activities Can Be Represented Using the Language of Maths

In Chapter 2, we saw how philosophers and researchers developed theoretical structures and principles to represent intelligent thought.

Scientists later applied these to the concept of intelligent machines and created theoretical designs for such machines. Engineers and scientists brought these designs to life, using computers and electronics to create AI. The language used from early theory to modern practice was mathematics, and it remains the basis of how an intelligent activity is represented and processed in AI.

AI performs intelligent activity by manipulating mathematical representations to generate answers or improve results. The answers or results are also expressed mathematically, so AI may require ways of converting mathematical language back into a practical, useful output, such as turning up a radiator or telling a driver which way to turn.

Concept 6: AI Repeats Small Tasks Many Times with Different Data to Find the Right Result, Which Usually Feeds a Bigger Activity

A feature of most types of AI is that they perform calculations and tasks many times over, using different data or varying task details, to find and improve answers and results. The specifics of a problem determine the calculations, variations and definition of improvement. Eventually, some combination of them will provide the desired result. This will often then be used with the results of other tasks to support a larger activity. The more complex the intelligence being simulated, the more repetitions needed and/or the more complicated the maths in the step being repeated.

Illustration: Understanding AI Movie Recommendations Using the Six Concepts

These six concepts work together as a framework to describe how a piece of AI works, and can apply to any example. In the next section, we'll use them to go through the main types of AI in common use today. But first, we'll illustrate the approach with an everyday AI example.

Introduction to AI Movie Recommendations

If you use a video streaming service to watch movies, say Amazon Prime, Netflix or Disney+, you'll be familiar with the idea of personalized recommendations. This is where the service provider suggests movies it thinks you'll enjoy. The AI behind the suggestions is called a Recommendation Engine, and is a common use of ML.

These companies usually have many millions of customers and hold information about the movies each customer has watched. Customers can provide an online rating of a movie after watching it. They may also have provided some information about the kind of movies they like through surveys or a social media profile.

Applying Concept 1 to Movie Recommendations (A Well-defined Problem)

To apply Concept 1 to this example, we need to describe the problem we want to use AI to solve, in such a way that it's suitable for ANI, that is, well-defined with clear boundaries.

In this case, the movie provider wants to use AI to improve the way it recommends movies to each customer, so they're more likely to enjoy them than the current recommendations. If they do this well, each customer will watch more movies; there'll be less likelihood of customers leaving, and a higher chance of customers recommendations.

So, the intelligent activity to be automated and improved is: 'suggest a list of movies a customer will enjoy'. AI could be and is used in many ways by a streaming service, but the first concept focuses us on this specific, single one.

Applying Concept 2 to Movie Recommendations (Presence of Intelligence Characteristics)

To apply Concept 2 to this, we need to confirm the types of intelligence involved in recommending movies. If there's none, then this isn't an

AI problem. Once we know which types are involved, we can look at existing solutions and tools for those types of intelligence that might be useful here.

In this case, movie recommendations need reasoning and learning. We know this from a mix of experience and common sense, so it may not be immediately obvious. With experience, it becomes easier to judge. Knowing this provides insight into how AI is improving the activity and helps our understanding of what the AI technology is doing.

Applying Concept 3 to Movie Recommendations (Comprising Several Smaller Activities)

We won't be able to tell from any given set of movie recommendations how the company used AI to make them, or even if AI was used. Concept 3 means breaking down the activity of making a movie recommendation into simpler tasks, focusing on the intelligent ones. This helps us think through the problem in the right way to understand how AI could be used to solve it, and the intelligent building blocks required.

Let's look at how we might break down the task of figuring out what kind of movies someone likes. One option would be to do some research on the customer, such as finding out what movies they've seen before, and how much they liked each one. We could look for patterns, trends and associations that tell us what they like. This might work if we had lots of data about a customer, that is, they'd watched lots of movies, and we knew what they thought of them.

Another option is to find other customers similar to them and see what kind of movies the similar customers like. This has the advantage of lots more data but introduces a new problem of finding similar customers.

Once we break down the recommendation activity into such tasks, some new ways of using AI might arise. For example, if we analysed all the movies watched by our customer, we might find that they like thrillers most of the time. We might then consider analysing the most popular thrillers across all the customers and see what the most popular thrillers have in common. We could perhaps then find features to use when making recommendations for this customer.

By breaking the activity down in this kind of manner, we start to build an understanding of the kind of intelligent tasks needed at a lower level. In this case, they're all forms of data analysis, forecasting and prediction.

Applying Concept 4 to Movie Recommendations (Fuelled by Data)

Recommending movies needs data about customers and movies. The first will be whatever each customer shared when signing up for the service, plus anything they've provided since in surveys or feedback. In addition, the company will have details of movies each customer has watched, any ratings they've provided, and other information such as when watched and device watched on (laptop, phone, etc.). There's far more available data about a movie than a customer. Most won't be held or owned by the movie streaming firm, but there'll be plenty available to access and use.

Applying Concept 5 to Movie Recommendations (Represented Using Maths)

In Concept 5, we look at what we can do with the relevant activities and data by translating everything into appropriate maths language. The data about customers and movies isn't particularly complex, but the volumes are significant. Customer numbers for large streaming companies are in the millions, movies numbers are in the tens or hundreds of thousands, so numbers of movies watched is in the billions or tens of billions.

These three inputs (all customers, all available movies and movies watched by customers) are available to use in an AI algorithm to create the output we're seeking. That output is a list of movies that a given customer might like, and ideally how likely they are to enjoy each. One skill of Data Scientists is finding ways to create the desired output from the available inputs. This will require intermediate steps of calculations and analysis, in this case to identify details such as patterns among movies each customer liked and characteristics of customers who liked certain types of movies.

The calculations, analysis and other manipulations of data can be modelled and performed using statistics, maths and computing. This is possible if the activities and data are represented in mathematical language. AI techniques and tools perform calculations, analyses and data manipulation efficiently and quickly for large volumes of data.

Applying Concept 6 to Movie Recommendations (Repetition with Different Data)

The aim of this concept is to reinforce that for AI to find a desired result, many tiny experiments and trials need to take place, applying various mathematical models to different types of data. Those which generate better results will be kept and refined, while ones that don't will be rejected. This leads to a set of mathematical steps that are repeated many times on the input data to create many intermediate results, from which the desired output can be identified.

Suppose we wanted to make movie recommendations for a particular customer who happens to be a male student aged between 18 and 25. One small AI activity to repeat many times is to check every customer's details, and if they're male students aged between 18 and 25, list their top 10 favourite movies. This would give us an intermediate subset of movies to analyse further. We may be able to perform further steps to make recommendations from within this subset. Another AI step might start with analysis of the particular customer's viewing habits, telling us that he likes thrillers, especially Swedish thrillers. We could then analyse other characteristics of thrillers, and people who like them.

Analyses and steps such as these would all involve repeating small steps many times, creating intermediate lists of customers and movies, then checking and analysing them for patterns and exceptions.

Bringing It All Together

The six concepts described here illustrate how intelligent activity, mathematical representation and data work together in AI. What they also show is that AI is currently a complex engineering problem that depends on human judgement and expertise. The most difficult parts of creating AI are understanding how best to use the available data to generate the desired results, and designing the algorithms and other mathematical steps involved.

The six concepts have been summarised in Table 3.1 below, and the next section shows how they are applied to some of the most well-known forms of AI in common use today.

Table 3.1 Six Key Concepts behind Narrow AI

Concept 1	AI needs well-defined problems, with specific boundaries.
Concept 2	Intelligence means the presence of at least one of the eight characteristics.
Concept 3	Intelligent activities are usually comprised of several smaller ones, only some of which may be intelligent.
Concept 4	Data is the fuel of artificial intelligence and machine learning.
Concept 5	Intelligent activities can be represented using the language of mathematics.
Concept 6	AI repeats small tasks many times with different data to find the right result, which usually feeds a bigger activity.

DEMYSTIFYING SOME COMMON EXAMPLES OF AI

The previous section showed us a set of concepts which we can apply to any situation involving intelligence, and use to understand the principles of how AI would apply to that situation. We'll now turn to four major areas where AI is most commonly used today and put that into practice.

- Intelligent analytics
- Computer vision
- Natural language processing
- Intelligent automation

There are of course many more uses of AI, and more appear every passing week. We're looking at these four areas because they cover the majority of existing AI applications you're likely to come across.

For each area, I've described how this type of AI is used, then walked through how the concepts apply. This deconstructs how the AI works in principle, as far as is possible without delving into technical details.

Intelligent Analytics

Analytics[34] is a form of AI that is familiar to many and has been with us the longest. There are several terms used to describe it, none tightly defined or universally agreed upon. Common synonyms, some with other meanings beyond AI, include predictive analytics, predictive forecasting, big data, and data analytics. It refers to the ability to automatically inspect computerized data (typically numbers and text), intelligently manipulate it, find insights within it and present it.

The three types of intelligent analytics we'll consider are:

- Intelligent search
- Forecasting and prediction
- Anomaly detection

Intelligent Search

Intelligent search[35] is the use of AI to find answers to queries about data held in large quantities, often across multiple sources. The best-known examples are Internet search engines such as Google and Bing, but they also appear as search facilities on databases (e.g. employee records, product catalogues) and websites.

Intelligent search has come on in leaps and bounds since the early days of the Internet, and now relies on AI. One feature is the use of unstructured data. This means data that doesn't conform to the strict rules that used to apply to computer data before it could be stored and retrieved. Today, it needn't even be in traditional computer form. For example, it's no longer necessary to search databases using references like customer number. This is because AI allows search to perform fast enough using partial text in names, other characteristics or even free text such as descriptions.

Another powerful use of AI in search is the use of natural language processing (NLP). Natural language interprets the search terms and uses intelligence to guess the intent of the searcher. NLP also enables spoken inputs instead of typing.

Forecasting and Prediction

Forecasting and prediction[36] using AI are actually two different AI applications, but the terms are often used interchangeably, and some new definitions even contradict previous usage. We can treat them as one, because for our purposes, the underlying AI is similar at a high level, and only changes materially when getting into details.

Both use AI techniques to identify patterns in data and use different algorithms to guess values for future data points. The algorithms depend on existing data available, including the type of variations expected. This can also be enriched with data from third parties.

So, for example, when planning a new outlet, a fast food chain could have huge amounts of data from existing stores, as well as third-party data from the landlord and local authorities, covering factors that might affect store performance. Data Scientists can use them in complex models to analyse and predict store performance, including diverse factors like footfall, presence of competitors, and distance to local transport stops. By filling this model with data about a proposed new store and its locale, they can forecast customer numbers, order values, staff numbers and so on.

Anomaly Detection

Anomaly detection[37] is the use of AI to spot exceptions and oddities in patterns of data, flagging them for human or automated action if they require intervention. The best-known example is fraud detection. There are many other common applications, such as network monitoring and maintenance, including power, data and transport networks. The AI in anomaly detection is reasonably straightforward and is a variation on how forecasting and prediction is done.

To understand at a high level how all these work, let's look at how the six concepts apply to this kind of AI.

Applying the Six Concepts to Intelligent Analytics

Concept 1: AI Analytics Needs Well-defined Problems, with Specific Boundaries

AI Analytics situations are very problem-oriented, and invariably arise because someone is trying to answer a question such as 'how much profit will we make next year if we open 10 new stores instead of 5?'. So, each instance of analytics using AI will have its own well-defined problem with specific boundaries. Generically, they're all variations of the same activity: drawing insight from data, by identifying patterns, associations and exceptions. Whether the final outcome sought is a search result, the likelihood of fraud, estimated hours of sunshine or a bid price for an advert, they're all achieved using comparable steps and techniques.

Concept 2: Intelligence in AI Analytics Means the Presence of At Least One of Eight Characteristics

Again, this concept is straightforward to apply here. AI Analytics is all about Reasoning, which in practice means performing calculations based on logical conditions and comparisons. Spreadsheet users may be familiar with simple versions of these kinds of calculations and comparisons.

Concept 3: AI Analytics Activities are Usually Comprised of Several Smaller Tasks, Only Some of Which May Be Intelligent

The breakdown of activities involved in analytics is usually straightforward for humans to appreciate, because the activities invariably involve some form of arithmetic or mathematics to which we can relate. For example, fraud detection may involve finding patterns that represent typical card spending behaviour of a customer, and flagging exceptions. We may not be able to create those patterns ourselves, and certainly not with AI's speed and accuracy; but it's instinctively 'do-able', even if we can't personally do the calculations.

The same applies to virtually all forms of analytics AI, once you examine the underlying principles of how the forecasts, predictions, anomalies or search results are found. The maths may be terribly complex, but it's something we can relate to. It's unlikely to cause us to feel intimidated by the intelligence of a computer that can do it, just impressed by its prodigious speed and memory.

Concept 4: Data Is the Fuel of Analytics AI and ML

The data involved in this type of AI is understandable and straightforward, even if the volumes are incomprehensible. It consists of regular computer data, for example, customer and sales data for predictions, share prices for forecasting and product performance statistics to identify faults in machinery (anomalies). It can also include data not in the form of numbers and text, for example, Internet search using images instead of words.

AI analytics relies on far greater volumes of data than humans can comprehend. This usually means at least millions, often billions and more.

Concept 5: Analytics AI Activities Can Be Represented Using the Language of Maths

At the level this book addresses, representation of intelligent activity for AI analytics can be done using regular maths and statistics. In practice this means more advanced techniques like linear algebra built on 'normal' maths and statistics. In a more advanced treatment of the subject, we'd see that it's not quite as simple as that, but the extra complexity won't improve our understanding here.

Concept 6: Analytics AI Repeats Small Tasks Many Times with Different Data to Find the Right Result, Which Usually Feeds a Bigger Activity

Many of the tasks required in analytics don't need AI, because computing functions already exist to solve many small steps required. For example, there's no point in using AI techniques to calculate averages and medians, because computer languages can do that automatically.

However, where this changes significantly is in the application of analytics results to real-world situations, especially when human behaviour becomes a factor. So for example, a search engine may use AI to list the statistically best results first; but it's only by seeing which ones humans clicked on first that we can add the human interpretation of those results into the AI logic. Similarly, statistics may be enough to predict a likely fraudulent card transaction, but it can't reliably anticipate that a customer has decided to take an impromptu trip abroad, totally out of character with past spending and travel patterns. AI uses more sophisticated logic and data to start predicting and forecasting such real-world circumstances, something regular computing techniques would struggle to handle.

Computer Vision

Computer vision[38] is the use of AI with digital photography (imaging), so that computer systems can 'recognize' the contents of images, and 'understand' their meaning and implications.

The examples we'll consider are as follows:

- Image processing and recognition
- Text and handwriting recognition
- Video processing

The major barrier to computer vision used to be the difficulty in capturing an image digitally using image sensors, and the quality of the captured digital image. Originally, sensors appeared in desktop-sized scanners, and could only handle black and white (i.e. no grey) images such as text. Now, fingertip-sized image sensors in phones, laptops and household devices can create detailed colour images of quality previously only possible with film cameras. Once we have a high-resolution image in digital form, the role of AI is to 'understand' its contents. Before looking at how it does that, we'll explore three common uses of computer vision.

Image Processing and Recognition

Image processing[39] is the manipulation of digital images to improve or change them to support some wider goal. For example, Adobe Photoshop is a well-known tool to enhance photos, often used in magazines to make people appear more as they, or their editors, would like (them) to look, for example, slimmer or less wrinkled. This 'touching up' of digital images is AI image processing, using AI to identify and change parts of an image.

Image recognition is the mechanism by which an image being processed can be compared with other known images and matched with similar ones. The obvious examples are facial recognition in security systems and number plate reading in traffic control or automatic toll charging.

Text and Handwriting Recognition

Text and handwriting recognition[40] are specialized forms of image recognition. Text is easier, but they're both solved in a similar way. They rely on a concept called 'deep learning', in which different 'layers' of the AI handle different parts of the problem. The first layer identifies lines and shapes, the second compares of these with known shapes (i.e. letters, numbers and punctuation) and further layers perform further tasks on the letters.

Handwriting recognition is similar, with two additional difficulties. First, hand-written shapes are more varied and may be joined together, so it's trickier to identify where letter boundaries lie. Second, there's far more variation in the set of 'correct' answers to compare with, because there are many more variations in written versions of a letter than printed ones. Nevertheless, these issues can be overcome with more samples of handwriting to compare with, and more sophisticated algorithms for the more difficult detection.

Video Processing

For most applications such as home security and traffic monitoring, computer vision of video[41] works in a similar way to still images. AI used in movies for CGI and 'green screen' special effects are much more complex versions of image manipulation, but in principle use the same kind of technology and principles.

The major difference between video and still computer vision is the constant stream of images in video. This leads to an additional task of comparing images over time to detect changes, for example, traffic accidents. Movement towards and away from cameras can also be inferred, adding 3D to the potential usage.

This means there is far more data to process, and a new type of processing, comparing images over time. This processing is a variation on that used to analyse still images, but different AI techniques are needed to understand what any change means. Meaning can only be inferred from changes if there's an expectation to what should have happened. Hence, prediction tools and techniques are applied to video data, so that 'normal' behaviour and exceptions can be recognized.

Applying the Six Concepts to Computer Vision

Concept 1: Computer Vision Needs Well-defined Problems, with Specific Boundaries

We can't yet create general purpose computer vision devices that can switch between reading text, scanning crowds for suspicious behaviour, identifying known faces and appreciating art. So, computer vision applications need to be designed around individual problems that can be addressed with one type of computer vision. Examples include processing still images, adding special effects to video, reading handwriting, recognizing faces and identifying landmarks.

Concept 2: Computer Vision Intelligence Means the Presence of At Least One of Eight Characteristics

Computer vision AI is predominantly about automating perception, performing the intelligent activity that the human eye and associated parts of the brain handle. It's an AI replacement for one of the five human senses, sight.

For video images, there may also be analytics, in order to predict expected changes in images (such as the normal path of a car along a road) and anomalies (such as a collision with a lamp post). Image processing may also use this.

Concept 3: Computer Vision Activities Are Usually Comprised of Several Smaller Tasks, Only Some of Which May Be Intelligent

There's potential confusion here in the distinction between the smaller steps required to enable computer vision to 'see' something, and the smaller steps involved in an activity which uses computer vision, such as steering a self-driving car. To understand computer vision, it's the former we need to decompose.

The first step in computer vision is purely mechanical and is the conversion of the light forming the image into electrical signals that represent it. This is done by the retina in the human eye, by light-sensitive film in analogue cameras, and an image sensor in digital cameras. The electrical signal generated holds information about the image, including brightness and colour of each part of it.

The rest of the tasks in computer vision are all intelligent, and involve extracting meaning from the electrical signal, and deciding how to process it to perform an activity, for example, surveillance or document scanning. Depending on the activity, this may involve manipulating the signal to change its meaning or sending a mathematical description of the image to another task, such as displaying it on a screen or changing a steering wheel direction.

'Manipulating the signal to change its meaning' could be something innocuous like removing a blemish or sharpening a blurry detail. But it could also be something more dramatic, such as in so-called deepfakes.

Concept 4: Data Is the Fuel of Computer Vision

The data involved in computer vision is the representation of an image as a series of tiny dots, usually not visible to the naked eye. There are typically one or two million such dots on a laptop or phone screen. These dots are known as pixels, and each has a colour which can be defined using numbers.[42] For images on screens, the colour is expressed as a set of ratios of red to blue to green.

In terms of data volumes for computer vision, the two factors to understand are numbers of pixels in an image, and numbers of images. Together these provide a sense of how much data needs to be manipulated, processed and moved around between tasks and devices. Phone camera images are representative of computer vision images and have sensors up to a few tens of Megapixels. This means the images are tens of millions of pixels in size. Each pixel consists of three numbers (red, blue, green), so a single typical image may have tens or hundreds of millions of data items. Computer vision problems such as facial recognition, visual search and video monitoring typically involve at least a few thousand images, but more likely millions or tens of millions, especially for video.

So, the overall amount of data involved in computer vision quickly mounts into hundreds of billions or more (tens of millions of images, each containing tens of millions of data points). These typically need to be processed instantly or in milliseconds. Hence, the research emphasis on computing efficiency and speed.

Concept 5: Computer Vision Activities Can Be Represented Using the Language of Maths

The representation of images in mathematical form is deceptively simple in principle, because an image consists of dots, and each dot consists of three numbers (three for colour). So, an image can be represented as just a string of numbers. For example, if an image is made up of 100 dots in a 10×10 grid, then it can be represented by a list of 300 numbers.

The manipulation of images also appears to be relatively straightforward conceptually, because it's possible to describe visual features as mathematical features in the strings of numbers and make image changes by applying mathematical operations on the numbers. For example, the edge of a black object on a white background will be a row of pixels each with brightness of one (white) right next to a row with brightness of zero. So, finding edges on an image requires searching for sets of numbers corresponding to pixels that match those patterns. Similarly, to change a colour on an image, the string of numbers is searched for the red/blue/green value of the colour, and each occurrence is replaced by the red/blue/green values of the new colour.

However, while both consist of easy to understand concepts, they are impossible in practice without advanced, sophisticated mathematics, that allows the principles to be applied to the huge volumes of complex data in meaningful timescales. Hence, many technical innovations in AI have come from computer vision work.

Concept 6: Computer Vision Repeats Small Tasks Many Times with Different Data to Find the Right Result, Which Usually Feeds a Bigger Activity

From the description of what image data is, and how it can be searched and transformed, it might be clear that computer vision is about repeating such searches and transformations many times to achieve an overall result.

For example, facial recognition is about matching an image of a face with a set of known images in a database. It is hopefully clear from the walkthrough so far that there are two broad but computer-intensive steps. First, a face image needs to be converted into a mathematical representation. In this case, using strings of numbers to represent pixels isn't efficient enough, so it happens to be a different mathematical technique. The second step is to compare the maths description of the pixels of this face with a similar representation of other faces in a database. Unless the lighting and angle are the same, it won't be a match. So, the algorithms have to allow for such differences. But in principle, the computer vision is comparing two massive strings of numbers, representing two sets of pixels.

We saw that each image may consist of hundreds of millions of such numbers, so the comparison of one image against a database of hundreds or thousands is not to be underestimated. The only feasible way to do this today is to break up the images and strings of numbers into thousands or millions of pieces, comparing each piece at a time, then reconstituting the pieces. The pieces could first be individual pixels or small groups of pixels, then bigger groups such as an eye or nose, and eventually a whole face. There are mathematical shortcuts using something more efficient than comparing pixels, but these are the kind of steps involved, hence the need for powerful computing for computer vision.

One of the many techniques that reduces images into manageable-sized elements is known as the 'sliding window'. This involves only examining a small rectangular window of a few pixels in size, and using AI to identify its small contents, say a short diagonal line. The window then slides across the image to inspect an adjacent set of pixels, which may overlap with the first, and identify again.

In this way, the AI builds up a representation of the overall page in pieces that are easier to compare, manipulate and use. This is a typical example of a small task that is repeated many times with different data.

Computer vision is full of many such steps that together analyse, recognize and manipulate images.

Natural Language Processing

NLP[43] consists of three broad types of AI activity:

- Natural language generation (text and speech)
- Natural language recognition (text and speech)
- Natural language sentiment analysis

These might sound similar, but the differences lead to very different challenges and levels of difficulty.

Natural Language Generation

Generating natural language, especially text rather than speech, is the least complex form of NLP. The most rudimentary form isn't even AI, but simple programming (coding error messages and system messages using natural language). The main way it's done is to use databases of vocabularies, sentence structures and synonyms to rewrite text into new, natural forms. AI's logic is used to derive the intention of the text, find and construct alternative ways of expressing that, and evaluate the best one to use. Smart speakers are the most obvious example, where there can be variety in how the same message is worded, even after several occurrences.

Once the choice of words and form of language have been selected, there's no AI required to present them on a screen. However, if they are to be spoken aloud, a different type of AI is required to convert text to speech.

AI in computer vision relies on images being converted into pixels. Words have no direct equivalent to the pixel, but AI makes use of an approximate audio equivalent called the phoneme.[44] This is a type of sound which makes up a spoken language, and is similar to the phonetic approach for teaching infants the alphabet and reading. AI speech generation converts text into phonemes as one of its tasks.

Natural Language Recognition

Recognizing natural language[45] is a complex problem, and at its most extreme, approaches general AI levels of difficulty. However, in practice AI today is only expected to recognize speech or text in defined circumstances (such as home automation or operating TV controls). Within the limitations of those circumstances, it's already impressive, and continues to improve.

The most difficult part of recognizing natural language is dealing with the diverse usage of language. This leads to many ways of saying the same thing, and many similar sounding words and phrases with very different meanings. The lack of consistent, universal rules to dictate language use only makes the difficulty greater. Finally, human communication contains both unspoken elements within language and non-verbal messages to complement language. For example, a phrase such as 'the teacher beat the student on the athletics track' is technically ambiguous, but most people would assume it's just clumsy phrasing, and it's describing a student–teacher race rather than public punishment. But it's impossible to train AI to consistently deal with all such examples.

Natural Language Sentiment Analysis

Sentiment analysis[46] is the ability of AI to understand the sentiment and emotion in a piece of written or spoken language, which adds a further level of difficulty to language recognition. Its use is widespread and varied, but effective in relatively few areas so far.

A common one is automatically reading reviews and comments about a product or service, and understanding the feelings expressed by customers. It's especially useful in scanning social media posts. However, this isn't sentiment analysis the way humans recognize it, but more of a statistical exercise, closer to AI forecasting. It's done by finding particular words associated with certain emotions and looking for their frequency of use. It's more complicated than that in practice, but it's based on using the frequency of certain words as a proxy for sentiment.

Humans find the sentiment and emotion in a piece of language in a different way, reading sentiment in the individual words in a sentence, and AI is still early in its journey to replicating that accurately and widely. Currently, recognizing subtle language features such as irony, sarcasm and even dry humour is beyond AI. And of course, NLP is restricted to language, whereas much human communication, especially of sentiment, is non-verbal.

Applying the Six Concepts to NLP

Concept 1: NLP Needs Well-defined Problems, with Specific Boundaries

Generating natural language is a pretty narrowly defined problem statement, and doesn't create much room for ambiguity. The scope of the problem will be defined by the language itself (English, French, Hindi, etc.) and perhaps the area of application (general use, medical diagnosis, sports commentating, etc.).

In contrast, the main challenge with NLP recognition, especially speech, is the breadth of possible language used. There's near-infinite variety of language usage, form and presentation (e.g. accents, idiom) in normal human conversation. Without any constraints on these, true NLP voice recognition is close to general AI, in that it needs to be ready to cope with almost any possible meaning. The more constraints that can be applied, the narrower the scope of language that can be expected, and the better the performance.

Concept 2: NLP Intelligence Means the Presence of At Least One of Eight Characteristics

Natural language communication is one of our eight characteristics of narrow intelligence. But perhaps surprisingly, NLP AI doesn't only need this type of intelligence. That's because while this a single type of intelligence in humans, in computers it's not. Natural language communication in humans is a sophisticated, complex capability that is not yet fully understood. Our ability to simulate it artificially has reached different levels in different areas. For some aspects, we still need to supplement it with other forms of intelligence, specifically AI analytics.

Conveying messages using text or speech is something computers can do easily, and making it sound natural can also be achieved pretty well by referencing dictionaries and sample text. But the step to understanding human language is significant and can only be achieved by adding significant amounts of AI analytics. There's also technical difficulty in understanding context, in that a word's meaning may change depending on the words before and after it. So, AI may need to remember whole sentences and phrases to accurately understand individual words within them. Similarly, sentiment analysis is still in its early days, and requires complex processing to identify the many ways emotion can be conveyed through words.

Concept 3: NLP Activities Are Usually Comprised of Several Smaller Ones, Only Some of Which May Be Intelligent

Making computers output text onto a screen is a basic computing task, and with minimal effort, can simulate natural language generation. For example, a program designed to tell the time could be programmed to output 'Hello, the time is' followed by a computer instruction to display the time. This is the most basic form of natural language generation, albeit not very intelligent. To make this AI, we could break this into smaller tasks of choosing more varied words, and flexibly handling different situations. This would be a combination of logic to determine the circumstances in which the given words could be used, and the data about other forms of language that could apply to those circumstances.

In any form of natural language speech generation, the words required are generated in text form in the same way. AI speech generation then adds an extra step to convert that text into phonemes that will sound like someone reading it aloud and playing them through a speaker.

Understanding natural language also splits into spoken and textual. As with speech generation, dealing with spoken input adds an additional step to convert spoken to written text, then follows the same tasks. The tasks to recognize and extract meaning from natural language are incredibly difficult, and require complex algorithms, analysis and processing. This is because of the vast number of ways language can be used, and the inherent ambiguity in everyday language use.

Concept 4: Data Is the Fuel of NLP

The data used by NLP consists of words, phrases and combinations of them, all comprised of text. It also needs data about language structure, usage and meaning.

While there are a lot of words in a language for the human brain to remember, a typical dictionary is not large for a computer. So, vocabulary data is relatively simple and limited in volume, generally no more than hundreds of thousands of items. Data about language usage is much more complex and includes huge volumes of word combinations. The processing involved in language usage is much more intensive than that required for handling vocabulary.

Unlike the other AI applications so far, this type of data is flexible and fluid, and doesn't conform to rigid structures and rules. This makes language usage data difficult to categorize, process and use. To add to the complexity, much communication between people is non-verbal, so language by itself may not even be sufficient. Non-verbal communication cannot currently be reliably converted into data for AI to use.

Concept 5: NLP Activities Can Be Represented Using the Language of Maths

Representing language in maths is both easy and difficult. The easy part is converting letters, words and sentences into some kind of structured data. Children write secret codes to each other by converting letters to numbers, breaking them into groups or performing simple arithmetic on the numbers. Converting language into mathematical form for AI is a very advanced version of this.

The difficult piece is representing meaning, context and even sentence structure. There are all sorts of techniques to help with this, and collectively they can go a long way towards the goal, especially for common language usage. But completing the final steps and dealing with large numbers of exceptions is still hugely challenging.

As a result, generating mathematical representations of natural language is difficult but manageable. Limitations in data, algorithm complexity or processing power, don't usually prevent natural language being created, they only restrict its possible richness, variety and sophistication. But that's not the case with natural language recognition, which needs representation that's sufficiently comprehensive to deal with any possible natural language input. This reinforces the very first principle of the narrowness of the problem. If the type and variety of natural language to be recognized is sufficiently narrowly defined, say just for conversations about medical symptoms, better results are possible.

Concept 6: NLP Repeats Small Tasks Many Times with Different Data to Find the Right Result, Which Usually Feeds a Bigger Activity

There are not the same levels of repetition involved in text and speech generation as other types of AI. Some may be required to fine tune how natural different intonations sound or choose the most appropriate word from a selection. There will also be a degree of repetition of tasks to train such applications to sound natural from the outset.

Speech recognition is at the other extreme. That's because the detailed steps involved in preparing speech for recognition are intensive and create many paths and possibilities of what some words might mean. Each of these needs to be processed, evaluated and compared with other possibilities, before moving onto the next word.

Examples of the preparation[47] include stemming and lemmatisation (removing different endings from words to simplify), parsing (grammatical analysis) and topic modelling (uncovering hidden structures in large quantities of text by inspecting distribution of words).

Intelligent Automation

Intelligent automation[48] is the use of intelligence to automatically follow processes based on conditions, inputs and instructions. The difference between regular and intelligent automation is the ability to deal with flexible, vague or unexpected circumstances. The type of process being automated can take many forms, from processing information such as orders, through changing settings on devices such as home entertainment systems to operating machinery such as electricity generating equipment.

Intelligent automation can also involve other intelligent features. One is the use of NLP to receive instructions and output results, for example when automating customer service calls. A second is the use of analytics to assess the circumstances and decide on an action, for example when automating the acceptance or decline of a loan application. The third is the use of motion and manipulation to physically control machinery, such as moving products in a warehouse to a loading bay to be despatched.

At a very high level, the examples we'll consider all work in similar ways, even though the areas seem very different. We'll see why that's the case when we walk through the six concepts to deconstruct how intelligent automation works. Here are the four types we'll explore.

- Virtual agents/robotic process automation (RPA)
- Internet of Things (IoT)
- Robotics
- Autonomous vehicles

Virtual Agents/Robotic Process Automation

RPA[49] and virtual agents are used to automate what used to typically be paper processes such as sales orders, accounting ledger entries and insurance claims. These are usually performed using computer systems today with paper copies sometimes only printed for compliance, legal or customer preference reasons.

AI is used in such processes to automate the work people would do to operate the computers, such as entering customer details into a computerized order processing system or sending a quotation to a customer. The human agent is replaced by an artificial agent to automate the process. The artificial, or virtual agent performs the process instead of using AI, hence the two possible names.

Internet of Things

The IoT[50] refers to the addition of sensors on all sorts of devices, so that the information about the device can be sent to and from a computer system. The computer system may do something based on the status of the device, such as turning up a radiator if the temperature sensor in a room falls below a desired level. It may also control the device itself, for example adjusting the timing of traffic lights depending on the volume of traffic.

This data is transmitted via wireless or wired networks, and often over the Internet. As the Internet is being used to connect devices rather than computers, this is described as the IoT, the 'things' being the connected devices.

The intelligence involved is in how the data from the devices is processed, and decisions made about the actions to take. The simplest decisions and actions require no intelligence and are based on a set of pre-defined rules contained in a computer program. When the decisions get more complex or ambiguous, especially if there are several valid options that need to be evaluated and compared, intelligence is required. These then move away from using conventional computer programs, and instead use AI algorithms to make the decisions.

Robotics

Robotics[51] is the presence of intelligence in mechanical machinery that physically moves. The machinery may remain in one location, with the movement happening within the machine, such as an AI drill that uses different settings for different products and components. It may also involve the machine moving around under the control of AI, such as robot vacuum cleaners or the kind of robot butlers seen in science fiction.

The AI is used to decide what part of the machine should move, and how it should move. For example, how much to adjust the drill depth, how many metres the vacuum cleaner should move before starting to clean, or where a robot butler should take a glass of beer to serve it to a person.

Robotics doesn't just mean robots as we see them in the movies; it means intelligent automation that involves movement of any kind.

Autonomous Vehicles

Autonomous vehicles[52] are ones which are able to move, change direction and stop without human involvement. Self-driving cars are the most advanced examples. But factories also contain fork-lift trucks that carry loads to destinations, and there are many other industrial examples that are less glamorous than self-driving cars.

Autonomous vehicles are not just examples of intelligent automation. They are complete ecosystems of IoT sensors, robotics, computer vision and analytics.

Applying the Six Concepts to Intelligent Automation

Concept 1: Intelligent Automation Needs Well-defined Problems, with Specific Boundaries

Each example of intelligent automation described above covers a type of ANI problem or set of problems and can't be used to solve a different one without adaptation. For example, an insurance claim process needs to have the process defined clearly before RPA can be applied to it. That RPA solution cannot then be applied to posting ledger entries into an accounting system. Similarly, the AI controlling an industrial drill using data from IoT sensors can't simply be applied to a lathe on the same production line. There's a slight exception with robotics and autonomous vehicles, where the robot or vehicle can do several things that can be automated using AI. In those cases, each of those automated activities is a separate AI problem, with its own definition and scope. The overall robot or vehicle can therefore do several different automated activities using AI, each of which is narrow AI.

Concept 2: Intelligence in Automation Means the Presence of At Least One of Eight Characteristics

Intelligent automation uses reasoning to work out what to do and may use motion/manipulation if the actions involve movement. Others will be involved in autonomous vehicles and robotics, but not necessarily as a core activity.

Concept 3: Intelligent Automation Activities Are Usually Comprised of Several Smaller Ones, Only Some of Which May Be Intelligent

There is great diversity in practice between the different types of activities that can be automated using AI, but they can invariably be broken down into a series of conditions, decisions and actions. The AI principles behind this kind of intelligence are similar across applications, but change significantly in how they're applied.

There may be many levels of decomposition to break up a large process into steps in the form condition/decision/action, and these then need to be reconstituted into a larger overall activity. The example of making a humanoid robot walk breaks down into automatically operating several individual motors on its feet and legs, to create the overall activity of taking steps. As mentioned, there will probably be other activities involved as well, some intelligent, for example computer vision to make sure the robot isn't walking into an obstacle.

Concept 4: Data Is the Fuel of Intelligent Automation

The data involved in automation varies with the activity being automated but is usually obvious in each situation. For example, automatically posting an order into a sales ledger requires data items on the invoice and data about posting rules. Meanwhile, automatically operating brakes on a car requires data about the vehicle, its surroundings and rules for braking (vehicle speed data, visual data about the presence of other traffic and pedestrians, mathematical data about braking distances and so on).

Data involving IoT, robotics and vehicle automation tend to involve the greatest volumes and complexities. RPA can also involve huge volumes, but these are typically simpler in form, and the processing involved is less complex. The data volumes relate to the number of transactions being processed, number of IoT sensors and frequency of feeding back device data or vehicle driving decisions required. This causes challenges, but doesn't generally present as many difficulties as the other types of automation.

Concept 5: Intelligent Automation Activities Can Be Represented Using the Language of Maths

The maths used to represent RPA processes is relatively simple compared to many other types of AI. Simple flowcharts can be sufficient to describe processes, even complex ones, and are straightforward to represent in mathematical form and computer programs. When a process has many conditions to test and options to choose from, the flowchart describing it may be large, but the maths involved does not get much more challenging, and there's just more of it. For the most complex processes, advanced maths and statistical techniques may be required to cope with the very large numbers of permutations of decisions, paths and outcomes.

In IoT automation, the outputs and decisions in the processes relate to devices, and this is also not complicated to represent. Flowcharts still broadly work to represent the process being automated, but with additional notation to represent devices. This in turn leads to different forms of maths. For more complex sensors and devices, there may well be more complex maths required to handle the signals being processed, but this still doesn't present major challenges usually. And as with RPA, for very complex processes, there are additional maths and statistical techniques required to deal with the very large numbers of combinations to consider.

Robotics and autonomous vehicles both involve movement that is controlled by AI. Movement is represented by coordinates and arrows between them, which can be converted into mathematical notation relatively simply. A difficulty arises when there is lots of movement happening, being calculated and being instructed. The difficulty is the size and number of calculations that need to be made nearly instantly to control the movement. As with RPA and IoT, advanced techniques from maths and statistics are used to help perform these calculations efficiently enough to be useful. For example, once a likely impending collision has been identified, the calculations required include assessing if emergency braking would prevent it, or whether changing direction sharply would be more effective.

Concept 6: Intelligent Automation Repeats Small Tasks Many Times with Different Data to Find the Right Result, Which Usually Feeds a Bigger Activity

With RPA, there is not the same emphasis as other forms of AI on finding ‘right’ and ‘wrong’ answers from many different potential answers. This is because most typical RPA processes are built on sets of logical rules, so finding ‘best’ answers doesn’t arise as often or in the same way. Where it does occur is in making calculations required by RPA process, but these usually involve a different form of AI, typically analytics or prediction. An example is using AI analytics to calculate an insurance premium as part of automatically executing an insurance quotation process.

With IoT, robotics and autonomous vehicles, the intelligent automation will typically be more complex, because the range of inputs to handle gets progressively wider in the three areas, and potentially more ambiguous. As complexity of circumstances increases, determining the ‘right’ and ‘wrong’ thing to do gets more difficult. As this happens, AI techniques that involve repeatedly considering and evaluating the best of many possible choices become more important.

HOW WE TEACH MACHINES TO 'LEARN'

The last type of machine intelligence we'll describe is ML, the ability for a machine to improve what it does over time based on past results.

ML is achieved using the AI equivalent of how children used to be taught at school: 'Practice makes perfect'. It involves a computer repeating a piece of intelligent activity many times, adjusting the activity each time until the results improve. Different types of ML use different types of adjustment and evaluation. At a high level, all types of ML follow a common approach, which we'll now look at. As we've done previously, our starting point is looking at the human equivalent of the artificial version.

How Humans Learn to Improve How They Do Something

Rather than exploring education theory or neuroscience, we'll stay with common sense and everyday experience. Usually, when we perform a task for the first time, whether learning multiplication or cooking, we know if we've got the right result. The teacher corrects or praises us, the meal tastes good or bad. So, a crucial part of learning is feedback on how well we're doing.

ML needs an equivalent feedback mechanism, some way for a computer to know which of its attempts to perform its designed activity are correct. It also needs to know how close wrong efforts have been. For example, NLP speech recognition will only get better through ML if the smart speaker has a way of knowing which words it understood correctly, and how close the wrong ones were.

Now let's look at what a human does with that feedback. The main task is deciding what to change to improve the result. This is the crux of ML and explains its dependency on data: Machines learn by analysing huge quantities of data about their own performance.

In most situations involving human learning, we use feedback on incorrect results to figure out what to do differently. If it's a multiplication error, we try harder to remember the table we got wrong. If it's a meal that tastes bad, we try to understand what was bad and how bad, such as if it was overcooked or under-salted. We assess a variety of factors to isolate the specific improvement needed, then try again with a change that should address it. Assessing the right factors, and even knowing what factors to consider, depends on our understanding of how the activity works, and our experience of improving it. Hence, teachers and coaches can help us learn faster.

How Machines Learn to Improve How They Do Something

For a machine to learn, it needs an equivalent strategy to get feedback, assess correctness and evaluate options for improvement. Once it has feedback on how well an attempt at something worked, it starts learning by repeating each step of the activity in as many different ways as possible. It evaluates the likely results of every possible next step it could take differently, churning through all possible next steps until it finds one that would give the best answer.

For example, a ML approach to recognizing a letter on a page might be to compare it with every letter in the alphabet, and set a score each time for how closely it matches. The score would be a mathematical calculation by the AI, which compares the maths representation of the target image with a similar representation of each letter in the alphabet.

Unlike humans, a computer doesn't initially 'know' which available steps are likely to succeed, so it has to try all of them. It rejects those that give a worse result and keeps track of ones that give a better result. This is why data matters so much in ML. If AI is seeking the best way to change how it does each step of an activity to improve it, it needs data for each different attempt of each step. Once it finds something that seems to be an improvement, it needs to repeat that enough times to be sure it's really an improvement, not just a fluke or coincidence.

Random Trial and Error as a Learning Strategy

A slightly desperate approach for a human to learn might be random trial and error. This isn't usually how people try to improve something they're not doing very well. But if it's something we don't really understand, we might resort to trying out random changes.

This rarely works other than by luck for humans and is usually a frustrating exercise. However, computers don't feel frustration, and can perform many tasks unimaginably quick, so random trial and error can be viable. The OCR example would theoretically be a good candidate for this, because the learning is simply comparing images of every letter on the page with 26 initial reference images of letters of the alphabet. With each correct guess, the computer will have another example of what each letter looks like, say in different fonts and styles. This will increase the number of reference images to compare with, and so increase the AI equivalent of confidence in future matches. In other words, it will learn to recognize letters better.

Better Learning Strategies than Random Trial and Error

Trial and error works as a learning strategy if we know whether each attempt is better or worse than earlier ones and can then select the one which was best of all. It's very inefficient, because we won't know the best choice available until we've evaluated all of them.

To speed up learning, there are mathematical techniques to reduce the number of guesses required to find improvement. These lead to the three common ways to design ML systems, each with its own name: supervised, unsupervised and reinforcement. Each works well for particular types of problem and data. There's also a fourth version called transfer learning, which is a variation of these.

Supervised Learning

Supervised learning[53] is the ML equivalent of having a teacher marking each of your answers as you work, so you can continuously adjust your remaining work to improve your marks.

If a human student had this, they'd check each answer as they go, to make sure they're approaching things correctly. But soon, they'd only refer to answers for questions they're uncertain about. In other words, once they had confidence in their understanding of solving each type of question, they'd not spend much further time checking that type. But if that confidence reduces, say for a new type of problem, they'd refer to the answers again until they had figured out how to solve the new type of problem confidently.

The ML equivalent involves giving AI a known set of data along with the correct results of performing an activity with it. This is known as training data and should be the representative of the kind of data that will be used in real situations. The AI performs the activity as designed on the training data, adjusting and refining how it performs the individual steps until it gets an acceptable percentage of them correct. It will then use these improved steps on any new data. The AI has now been 'trained', and works in a better way than it was originally designed, for data comparable with its training.

Illustration: Supervised Learning in Traffic Cameras

To take this from the abstract to the concrete, we'll use an example from computer vision. Not OCR this time, but traffic cameras. As usual, we'll break the activity down into steps, pick out a smaller task that requires intelligence (in this case computer vision), and look at how ML could be used for that piece of the overall AI.

The intelligent human activity to be performed by AI in this case is monitoring a traffic camera, spotting accidents, traffic jams or other incidents, and taking appropriate action. This might mean informing the police or changing the speed limit elsewhere to reduce the flow of traffic into the jam.

Many steps are needed, several of which are intelligent. For example, distinguishing vehicles from pedestrians; distinguishing between two stationary cars in a traffic jam and two that have collided; or being able to detect when rush-hour traffic has changed from acceptably busy to abnormal volumes requiring intervention.

The supervised learning example is the first of these: spotting the cars in an image. Easy and obvious for a human, but surprisingly difficult for a computer. The way it's done is to start by creating some logic that instructs the traffic camera system how to spot a car in an image. This logic is part of what's called the traffic camera's AI algorithm. Supervised ML is then used to train the traffic camera system, by showing it many real pictures of cars, so that it can adjust and improve the algorithm and other computer vision steps. The training ends when the cameras can recognize cars in the training data with enough accuracy for real-life traffic management.

The key to this learning is the training data: a set of images of real cars on real roads. The crucial feature of training data is that it needs to have the correct answers, in this case car images, identified and labelled. That's usually in the form of boxes drawn around each car in the image, labelled with the word 'car', and perhaps its make and model. Other objects such as people and lampposts might be similarly identified. (It makes no material difference to the example if the images are stills or video).

One part of the algorithm would describe what generic cars look like, including features such as shapes and sizes, presence of wheels, windows and so on. As we know, the language used for that description is maths rather than English. The AI can compare this mathematical description of a generic car, with an equivalent mathematical description of any other image, and quantify how closely they match. In other words, it can give itself a 'score' of how likely any image in the training data will be a car.

Supervised learning happens during training, when the traffic camera system compares its calculated score of whether something is a car with the label that tells it whether it was a car. Where it gets answers wrong, it will make adjustments to its mathematical description of a car or how it compares descriptions to improve accuracy. The skill of data scientists is in the sophistication and choice of ML techniques used to make these adjustments. The details of such adjustments aren't relevant here. What matters is that by the end of the training, the use of labelled images of cars has improved the original description of a car and how it's compared with new images, so that the system spots cars more reliably.

In case you think this is about identifying different models of car, it's actually something far trickier. A challenge for the AI designer in this example is dealing with incomplete images, such as seeing cars from different angles, or only seeing part of a car because of obstacles in front of it.

As with most of the examples used, this description of supervised learning has been simplified for clarity, to the point where it's no longer strictly accurate. This is a conscious decision, so that the concepts and ideas can be illustrated more accessibly.

Unsupervised Learning

We've looked at ML, and seen that it's essentially using feedback on the correctness of a result, based on large amounts of known data, to improve the way an intelligent activity is performed.

Unsupervised learning[54] is a version of this where feedback is not directly available, because suitable examples of correct results don't exist. This may be because we don't know what we're looking for, or it could be that there isn't enough training data available labelled with the 'correct' answers.

There are many practical examples in the first category: When we don't quite know what we're looking for, so can't use supervised learning because we don't know what a 'right' answer looks like. For example, if we had customers in an online store who bought clothes from us, and personal information about them such as demographics, employment and so on, we might want to know the characteristics of customers likely to buy specific products. We don't know what those characteristics are, so can't use supervised learning, as there's no way to label training data. Instead, we could use unsupervised learning to figure out what those characteristics are.

Unsupervised learning works by examining the data about the problem we're trying to solve, and looks for patterns, associations and exceptions to find possible answers to the questions we were trying to answer. So, with unsupervised learning, human intervention may be needed to apply sense to the possible answers.

In the e-commerce example above, let's suppose we wanted to find out what kind of customers buy red skirts, so we can promote red skirts to similar customers likely to be interested in them. The unsupervised learning system would churn through all the available data about all customers who've previously bought red skirts and look for patterns and associations. It might find that most of them are female, most have bought lots of other red clothes, and that they've also bought lots of skirts. So far, no real surprises. But it might also spot something unexpected, such as a disproportionately high number work in the travel industry, or have 'senior executive' in their job title, or live in Liverpool and list soccer as one of their interests. There's no way of knowing immediately if these are insights or coincidences, and they certainly don't tell us what we should meaningfully infer.

A human could take those results, and either apply some judgement to the results, or do some further investigation, such as customer surveys. But the AI system doesn't have that option. What it can do instead is repeat the exercise with further data, either fresh training data to use during development, or new data obtained with use such as monthly purchases. Doing this, the unsupervised learning might find many new examples to confirm an earlier pattern, such as customers working in the travel industry are more likely to buy red skirts.

The point of the example is not whether these answers are plausible or useful, but to illustrate how ML can be used to generate objective answers to questions humans can't confidently answer.

Reinforcement Learning (Also Known as Semi-Supervised)

Reinforcement Learning[55] is a variation of unsupervised learning that is based on the 'carrot and stick' approach. It's different to the other two types of ML because it doesn't concentrate on whether results of an individual's activity are right or wrong, but on the overall result of a series of activities. It uses the computer equivalent of rewards and punishments to change the individual's choices made during a complete set of activities.

The goal is for the AI application to learn over time the most effective combination of activities to get the optimal overall result. It continuously adjusts individual and groups of steps to maintain or improve that bigger result. For example, if reinforcement learning were used to design AI to play chess, it would set overall game victory to be the optimal result. Intermediate results, that is, individual moves would be rewarded or punished based on whether they help lead to a win or a loss, not whether a specific move was good in the short term, for example taking a low-value piece instead of a high value one.

Rewards and punishments in this context means, rather boringly, simply a numerical indication of likely success of the whole game. Rewards increase it, punishments decrease it.

To evaluate each move in terms of the best overall result, the ML system uses a degree of trial and error. It tries many, perhaps all possible, variations of each individual step in the bigger activity (the overall game), and starts to favour the ones which give it the best overall results (victory).

Reinforcement learning isn't as well-known as the other forms of ML, but its applications are common. The most obvious, given the description above, is teaching AI to play games. If you've heard of game theory, then you might also correctly expect that reinforcement learning can help solve other problems that can be 'gamed'. Examples include stock trading, online advertising placement and pricing, and even medical research.

For now, we'll complete this section on ML by looking at how scientists have modelled advanced ML techniques based on the human brain. This brings us to two widely used AI terms: deep learning and neural networks. Here's what they mean and how they work.

Deep Learning

Deep learning[56] is a family of ML techniques that help find more accurate and sophisticated answers to ML questions. It can be used for supervised, unsupervised or semi-supervised models.

The key idea behind deep learning is breaking down the learning process into a series of steps and representing each learning step as a connected 'layer' of processing. Each layer works on a different piece of the overall problem and makes its answer available to the other layers. The overall result of the whole activity is obtained by combining the different answers from the different layers.

These layers are usually illustrated as physical layers in a diagram, but you should be clear that this is not a literal picture, and the layers are computer programs. Each layer is a set of rules and instructions to perform calculations on data, and the result of those calculations is the output of the layer. The reason it's called 'deep' learning is because there can be many layers involved. Eight to ten is common.

We can return to OCR for a deliberately over-simplified, hypothetical illustration of deep learning. OCR is done using deep learning, but the way it's done in practice today is more complex than the form described below.

In technical terms, AI performs OCR by detecting the text in an image (i.e. distinguishing between text, images, decorative elements like borders and other items like smudges), then identifying what it's detected (i.e. recognizing letters and words). We'll focus on the first piece, detecting the text.

As with all such examples, it's built on the idea that ML involves a set of rules and instructions to achieve a result, that it uses feedback to improve those rules and instructions, and that deep learning consists of several layers, each of which performs a small step of the overall ML activity.

So, to use deep learning to detect text, we start by breaking that activity down into small steps, in this case three, and use a separate deep learning layer to perform each step. The first layer (remember, a computer program) examines all the dark parts of the image (i.e. data representing the printed text and any other marks on the page), and determines where the boundaries are, using appropriate rules and instructions created when it was designed. It then needs to tell the next step (layer) where those boundaries are. To do this, it needs to represent that information in a form that the next layer's computer program can receive and process. As we know, this representation is done using the language of maths.

So, the output of the first layer is a set of mathematical data that is the input to the next layer, describing the locations of all pieces of darkness in the image, including boundaries.

The purpose of the second layer is to recognize the shapes of the dark pieces that the first layer identified. It does this using a different set of mathematical rules and instructions, and represents the answer using different mathematical language. The output from this layer is a mathematical description of shapes and lines, along with where they are (which is what the first layer provided).

This in turn feeds into a third deep learning layer, which could be designed to match the shapes to letters and numbers and start the exercise of ascribing meaning to the contents of the image. So, by the third layer, we have a mathematical representation of a set of letters, which we could either output as the result of the OCR activity, or process further as part of a more sophisticated OCR application. For example, the OCR designers might include other layers or intelligence to spell-check the identified text or speak it out loud.

What makes this a version of ML, rather than just AI, is that the logic and results of each layer will change over time, based on the accuracy of previous results. So, each layer in our hypothetical OCR would get more accurate or faster over time at detecting lines and identifying shapes, and the overall model would get better at identifying letters on pages.

We've covered as much as we need about how deep learning works. It's now time to close the section on ML with the final piece of common ML jargon: neural networks.

(Artificial) Neural Networks

Neural networks[57] are a set of computing techniques that crudely mimic how one part of the human brain and nervous system work. They use an artificial version of a biological object called a neuron, found in the brain.

The human neuron transmits information around the brain, and the artificial version does something similar. It's part of deep learning because artificial neurons are used to transmit information between and around deep learning layers. Because there are many connected layers, and many points on each layer, the connections can quickly become complex. Hence, the artificial neurons form a network of connections, giving rise to the name neural network, or, to be more accurate, artificial neural network.

There are over two dozen types of neural network used in ML, but it's sufficient to learn about three major ones: feedforward, recurrent and convolutional.

Feedforward Neural Networks (FFNs)

The descriptions and examples of machine and deep learning in this chapter have all been sequential, in that the step performed in each layer is always followed by the next. There's been no mention of information flowing back to a layer that's already completed a step.

These kind of single-direction neural networks are called feedforward[58] because the information being passed from layer to layer by artificial neurons only goes 'forward'. This is fine for many AI problems, but not all.

Recurrent Neural Networks (RNNs)

RNNs[59] involve inputs being passed backwards to layers that have already completed a task, or require the result of a layer to be used later on in the activity. They're needed for certain types of problem such as processing audio tracks or recognizing passages of language.

One question that arises about them is how the layers 'remember' the results of a step for future reference. If you're not technically minded, that might not sound difficult, but for many years it was. A major deep learning breakthrough was the concept of long short-term memory, first proposed in the late 1990s, which solved this.

The difference between FFN and RNN may sound a little academic, but there's a big practical difference. For example, if using ML for speech recognition, the meaning of a word will depend on the whole sentence in which it's used, especially if it has several meanings, or there are several words that sound the same. So, AI trying to recognize a word will need to know (i.e. remember) the words that have just been spoken, to decide what the correct meaning is likely to be. An FFN couldn't do that.

Convolutional Neural Networks (CNNs)

If you're not a mathematician or AI practitioner, CNNs[60] are a little tricky to understand, because the name comes from a particular type of mathematical operation called convolution. In simple terms, it's a way of intertwining two pieces of sets of data to create a more useful third form. It does this by applying a specific mathematical operation to the representation of the two data sources.

It's worth being aware of because CNNs transformed the field of computer vision. They allowed AI to recognize subtleties and sophistication in images that had previously been impossible. As a result, speed and accuracy of computer vision applications rose significantly.

4 TRANSFORMING BUSINESS AND SOCIETY

The world of the future will be an even more demanding struggle against the limitations of our intelligence, not a comfortable hammock in which we can lie down to be waited upon by our robot slaves.

—Norbert Wiener, Mathematician and Philosopher; Pioneer of Cybernetics

This chapter, as its name suggests, is about the impact of AI and ML on the world around us, both specific examples at home and work as well as its wider adoption by society. The things you'll read about are either already in commercial use or are expected to be very soon. In other words, out of the labs, being used and making money, or ready to start doing so soon.

We could have looked at healthcare, education, entertainment or many other areas. But we've left others to describe those, and in this chapter you'll read about AI encountered in the home, when out travelling, and at work.

IN AND AROUND THE HOME

AI in the home is about the activities to do with the home, such as heating and TV, and things that we usually do at home, but aren't about the home specifically. This includes tasks like online banking, booking holidays or listening to music. We'll concentrate on the first type here.

With all our examples of AI and ML so far, their creation involved examining an intelligent activity to improve it, breaking it down into smaller intelligent tasks, and designing ways for machines to do them better. To understand how AI has permeated the home, think about the things we do there in similar terms. It becomes clear that while there's a lot of AI around us, most of it is not the kind we'd find daunting.

Much of it is about automating tasks that require thought to do, but little that we couldn't do ourselves if we really wanted to. Examples include AI help with heating or cooking. We might do some of them more slowly than an AI version, or choose not to do them without intelligent help. But AI in the home today is primarily about convenience. The exceptions aren't really commercially mainstream yet, but will be technically feasible in limited form soon, for example, household robots. But let's start with something more familiar and comfortable to most of us.

Smart Speakers and Digital Assistants

The AI application that first comes to mind for most people is the smart speaker, or digital assistant. The best known of these are Amazon's Echo, Google's Home and Apple's HomePod. As explained earlier, these are a combination of electronic hardware, usually speaker equipped with microphones, and AI software. The software's main function is to understand spoken instructions and speak responses aloud. The software behind these devices is Amazon Alexa, Google Assistant or Apple Siri. Microsoft also has similar software, called Cortana, but doesn't produce its own smart speaker.

As with many AI applications and devices, their power lies in their existence within a wider AI ecosystem. This means there's a set of other devices and applications, usually connected wirelessly, which each do a single or small set of things, so that collectively the ecosystem has a wide set of functionalities. Examples include TVs and heating systems. These other devices and applications may use AI, but not necessarily.

With such ecosystems, one device typically coordinates the others, or is the main device for users to control and instruct other devices. With what's often called the smart, connected or digital home, the smart speaker commonly plays this role. So, while a smart speaker's nominal main job is to be a voice-controlled music player, most of its intelligence is used in its role as a 'hub' for other devices and applications, including ones using AI. Hence, an Amazon Echo or similar device could be called a smart speaker or digital assistant, depending on how it's used.

Controlling Other Household Devices and Applications

The other possible confusion about smart speakers is that the intelligent software used also appears in other devices, such as laptops, phones and TVs. This means that any device containing Alexa, Google Assistant, Siri or Cortana software could also be used as a digital assistant.

Common AI applications controlled by digital assistants include:

- Searching the Internet for information like travel and weather queries
- Performing simple online transactions such as banking, taxi booking and food delivery
- Organizing household tasks such as calendar reminders and updates, to-do lists and timers
- Managing heating, lighting, air-conditioning and curtains
- Operating household appliances such as vacuum cleaners, fridges and coffee makers
- Setting and checking home security systems and baby monitors

- Entertaining us with simple games, jokes and conversation
- Providing more complex entertainment such as finding and playing music, movies, TV programmes and games

Why Smart Speakers Aren't as Intelligent as They Sound

We saw earlier that most of the AI in a smart speaker is restricted to NLP, the ability to communicate with people about household activities using natural spoken language. The sophistication of this NLP is already very advanced, but continues to improve, along with the way it can be used.

For example, Amazon uses deep learning to improve the quality of speech, teaching its newest generation of devices to formulate full sentences before speaking them. Previously, speech was generated in words or syllables at a time, which created the more familiar mechanical-sounding voices we're used to, with slightly jerky intonation. Google has shown that its devices can speak with a personalized voice, rather than one of the standard ones. An early example used the voice of a celebrity—in this case, singer John Legend.

As well as using AI to improve the quality of artificial speech, smart speakers are also improving the way they are used. For example, intelligent doorbells are already available, consisting of a doorbell with a camera and speaker connected to the digital assistant. Previously, these haven't been particularly intelligent, simply sending video images of whoever's at the door to a mobile phone or laptop and allowing users to speak to them via the doorbell speaker and microphone. However, new versions of intelligent doorbells will be trained to have short conversations with visitors, using AI to understand what kind of visitors they are, and come up with different suggestions. For example, couriers could be asked to leave a package in a porch.

While this sounds very cool, and it is, from an AI perspective it's nothing special. The digital assistant is using NLP via the doorbell to communicate, and the kind of intelligence used in chatbots to decide what to say. A more advanced chatbot than the basic ones described earlier, but still nothing more than a set of rules on what to say in different circumstances.

All the other examples of intelligent applications beyond NLP controlled by a smart speaker reside outside the speaker itself and are only accessed by it.

Where the Intelligence in Household AI Lives

Most of the intelligent household applications listed above use basic AI for control and operation. The intelligence in most of the devices involves the following:

- *Perception:* Being aware of their state (e.g. on, off, loud), surroundings (e.g. warm, cold, people present), and instructions (e.g. turn on at 10 PM).
- *Reasoning:* Following rules depending on circumstances, for example, turning on, off, up and down at different times, temperatures or circumstances.
- *Natural language:* Hearing instructions and telling users when something noteworthy happens, like the burglar alarm going off.

These different types of intelligence work together to make the home more convenient, but the overall intelligence isn't very sophisticated. For example, intelligent heating systems mostly turn radiators up or down depending on instructions, ambient temperature, a schedule or an event (such as someone arriving home unexpectedly). The same applies to most things controlled in the home by a digital assistant.

For more advanced AI in the home, we need to look at how the digital assistant makes use of intelligence in applications it's connected to, usually through the Internet. These are typically services we could also access using our phone, laptop or tablet.

An obvious example is ordering a taxi by asking your smart speaker. The speaker will translate a request for a cab in 10 minutes to go the airport into an online order via a normal taxi ordering app or website. There's lots of AI going on to get the closest taxi to receive your instruction, confirm the job, find its way to you and then the airport, and estimate when it will collect you and reach your destination. It's just that none of it is in the smart speaker, or even your home. It's all being done by the intelligent applications at the taxi company or in the taxi. The only AI you've directly used is the NLP to tell your speaker what you want, and a little intelligence for the speaker to know which taxi company to use, and exactly what to ask for.

One big exception to this, one potential (but not yet common) use of very sophisticated AI within the home is robotics. As with other forms of AI, today's reality is the narrow version, and is already a multi-million-pound industry.

Robot Vacuum Cleaners

The best-known example of robotics currently available for the home is the robot vacuum cleaner. These are normal vacuum cleaners that use sensors and differing degrees of intelligence to clean rooms without human intervention or even presence. Similar technology has been applied to lawnmowers, to automatically cut grass in a garden and even spot weeds. Basic models aren't particularly effective, but the most advanced models are starting to use AI more effectively and improving with each iteration of development.

The way AI vacuum cleaners should work is a combination of:

- Motion and manipulation to move the device around a room;
- Perception to discover room size and shape, avoid obstacles and change settings for different flooring materials; and
- Planning or problem-solving to calculate the best way to clean every part of room without running out of battery charge.

They don't yet use AI as well as this, but will do soon,[61] and some manufacturers describe their top end machines as doing some version of this already. Let's look at what this will look like in practice, when the AI has been refined, and teething troubles ironed out.

The first time the robot cleaner is put into a new room, it will first teach itself about the room by moving around it, using sensors and a degree of trial and error, to find the walls, obstacles, carpets and so on. It will use this to create a digital map of the room, and then calculate how to cover the whole floor, how long that should take, and whether it will need to recharge its batteries to complete the job. It will set off to clean the room for the first time, covering each part of the room in line with its planned route. Once it's done this successfully for the first time, it will remember the room for the future, and use very simple ML to keep track of changes such as new obstacles. Manufacturers are working on more advanced AI, using computer vision, to identify obstacles and even dirt, so that if it comes across something wet and squishy, it will avoid it rather than spreading it around the room by trying to clean it.

This description sits somewhere between what's on sale now and what will be available soon. By the time you're reading this, that may already have changed. But hopefully what's clear from the example is that this isn't a very difficult AI problem to solve. It should be achieved sooner rather than later, through a combination of manufacturer development and hardware advances.

The much more interesting and challenging AI problem about robotics in the home is creating real-life versions of the kind of robot assistant we've seen for years in movies.

Robot 'Butlers'

We've read about human-like robots in fiction since Victorian times and earlier, but it wasn't until the 1960s that they became a familiar sight, brought to life on big and small screens. So we already have expectations for robots around the home, because we've seen so many fictional examples already.[62]

From C-3PO in *Star Wars* to J.A.R.V.I.S. in Iron Man, AI systems in mechanical human form have become well-known, including less well-intentioned versions such as in the Terminator movies. What makes these characters fictional, with no possibility of appearing in the real world soon, is that they exhibit complete AGI, and even superintelligence.

Every now and then we might see a news report about a new lab version of a robot that aspires to that kind of behaviour. But no matter how impressive they are (and some are incredibly impressive), they all fall woefully short of expectations Hollywood has set for AGI.

But if you instead start considering ANI in robotic form, things become rather different. If you've had beer with your meal at a Yo Sushi restaurant,[63] there's a chance it was served to you by a robot—a talking, motorized fridge on wheels. And if you've ordered snacks on room service at the Vdara Hotel in Las Vegas, robot dogs named Fetch and Jett[64] will probably have brought them to your room. These are robotic devices that intelligently perform specific activities and are available now. A robotic vacuum cleaner is a narrow AI robot, just not human shaped.

So AGI robot butlers in the home to do your bidding? Not anytime soon. But ANI robotic butlers to do specific activities automatically? Already happening, and creating new ones using current levels of AI and robotics technology isn't a huge technical challenge. They're possible for any household task that involves detecting, moving and simple manipulation. If the market is big enough to make money by selling it, someone will fund its production, and at least attempt to commercialize it.

The companies driving this area are an interesting mix, reflecting the different perspectives involved. Automotive firms like Honda[65] and Hyundai[66] are active in robot work, as are computer manufacturers like Asus.[67] Then there are the home automation and appliance specialists such as Samsung.[68] And finally, there are specialist robotics firms, like Hanson, creator of Sophia,[69] the first AI celebrity.

Firms like these are pioneering the use of AI in the home, working towards a much more sophisticated AI ecosystem. The results of their work will be smart homes that we control using much more intelligent devices than today. Depending on how accepting we are of change, it's likely that a home robot will be one of our options for tomorrow's version of the smart speaker.

OUT AND ABOUT

We'll look at two types of AI in this section, starting with how AI has become part of the vehicles and transport systems we use to get around, from cars and taxis to trains and planes. We'll also look at the kind of things we do while travelling, from figuring out where to go and what to do, to using AI to help us get there. We'll start with the AI innovation that has revolutionized the way we get around, navigation.

AI in Navigation Systems

One of the most familiar uses of AI, and a core technology in modern and self-driving cars, is satellite navigation. This appears in mobile phones and car entertainment systems, and there's still a market for dedicated navigation devices.

Initially, AI depended on an electronic chip in a device to receive signals from satellites circling the Earth, known as Global Positioning Satellites. These continually broadcast data about their position in the sky, which could be received by any device tuned to their frequency.

The AI in GPS chips would perform complex maths to calculate its position, by comparing the direction and strength of signals from the different satellites above it.

However, knowing where you are as longitude and latitude coordinates isn't helpful by itself for most people. You would previously have used the coordinates to find your position on a map, then used traditional map-reading skills to plot a route to where you wanted to go.

The big development in navigation AI that made it mainstream was the addition of digital maps. The first step was to digitize maps and add them into navigation systems. This allowed for new, more advanced navigation AI to do what people previously did: identify from the GPS coordinates where on the map it was, and the location of a user-defined destination. AI then performed the most complex piece of intelligence in the whole activity, that is, plotting a sensible route between the two points on the map.

How AI Plots Navigation Routes

Decomposing navigation into smaller steps leads to several intelligent tasks. Each of these is sophisticated compared to most household AI such as home automation.

First, the AI needs to represent a map in mathematical language, and place start and end points on it. The volume of data for just this is significant, as is the computer processing to perform the calculations involved. This increases substantially if the map includes real photos of the streets and buildings, which is now the norm.

Second, the most difficult part of the problem: calculating the route. Assuming you're using roads and paths rather than cross-country routes, this follows a similar principle to getting out of a maze. There's a basic, inefficient approach, which can be refined and optimized.

The basic approach is to start with trial and error, trying all possible paths from the starting position, and seeing which take you towards the destination. Eventually, if you tried all possible paths, one of them would be successful, but it would take a long time and many failed attempts. If you wanted the shortest route, you'd need to repeat this until you found every possible route and measure the travel distance for each.

To apply human intelligence to this approach, and rely less on trial and error, you would take each possible first step and look ahead for any obvious dead ends. You'd also look for long stretches that were clearly going in the wrong direction. You'd switch between a wider view of the whole map or sections of it, looking out for any obvious routes, and a closer, zoomed-in view of one part of the journey, to determine individual turns to try. You'd also be guessing the journey distance 'by eye', rejecting routes which were obviously longer.

The AI in navigation systems[70] does an electronic version of the same thing, using trial and error to find all available routes, wasting as little time as possible on obviously wrong ones, finally picking the shortest or fastest.

Digging in deeper, one way the navigation algorithm could work would be for the AI to simulate many imaginary physical footsteps on the map. After each footstep is taken (i.e. plotted), the AI would assess if this has taken it towards or away from the destination. This assessment would be a maths equation, comparing the distance before and after the step. The computer can recognize if a footstep is on a straight road with no turnings ahead, because straight roads and junctions are mathematical features on a digital map. If that's the case, it won't take lots of small footsteps, but instead take a single large step to the next junction, then assess whether it's closer or further away.

Doing this for every junction of every possible route is what turns this into a calculation involving millions of steps, in both meanings of the word. For obvious reasons, it can't stop evaluating a route, just because a turn on the way takes it further away from the destination.

With modern computer hardware and software, the AI is able to rapidly find every possible route on a map between two points. It does this through mathematical calculations that represent the map, the user's location and the steps along the way. The calculations include distance, so AI can pick the shortest.

Once AI was able to plot routes between points on a digital map, everything else we take for granted today involved developing variations and advances of the same principles. For example, finding places of interest along the way, calculating journey times (and therefore fastest instead of shortest routes) and adjusting a route to fill up with fuel or stop for coffee breaks. Each adds complexity and data volumes to the calculations but are relatively straightforward extensions. They're not simple mathematically, but neither do they rely on particularly difficult theory.

AI navigation is probably the most common use of AI in travel. Here's a run-through of some other important ones.

AI in Cars: Data Is the New Fuel

Former Intel CEO Brian Krzanich predicted that the average car will generate around 4 terabytes (4,000 Gigabytes) of data for every hour driven, and 10 times more for autonomous vehicles.[71] This compares with one or two Gigabytes of data produced by a person in a day through Internet use, video, messaging and so on.

Most of this will come from sensors in and on the vehicle, monitoring everything from images of the road to engine efficiency and tyre pressures. There will also be location and journey data, and of course cars will be Internet connected, so data will be coming and going via mobile networks.

He was speaking about the near future, but much of his prediction is already reality. If we look at a typical modern car using conventional fuel and needing a human driver, it likely already uses AI in a handful of ways.

Engine Efficiency and Vehicle Maintenance

AI uses data from an engine as inputs into algorithms that calculate adjustments to keep it running efficiently, for example the amount of fuel burned for different loads and speeds. The algorithms are designed using powerful computers to simulate real road conditions and identify types and amounts of adjustments. These simulations use ML to continuously improve, based on real-world and theoretical data. Improvements are turned into updated algorithms and control software, which can be uploaded to the car to make it more efficient. This is one of many reasons why a garage that deals with modern cars needs computers as much as it needs spanners.

A simple step from monitoring engine efficiency to monitoring the overall health of the car. Sensors can be placed throughout the mechanical and electrical components, giving a complete picture of how things are working and parts are wearing. Through this, modern cars now tell the driver when they require servicing, rather than waiting for a specific mileage or time period.

Safety Features

AI has been used for a few years already to improve passive safety through evolutionary developments. For example, how a car structure deforms during collisions,[72] or how airbags are deployed[73] for different drivers.

It's made more revolutionary changes through active safety features, where the car intervenes in some way to avoid or minimize danger. The most familiar example is the rear parking sensor. A car's rear bumper contains sensors that use a type of radar to monitor for obstacles when reversing. If one is detected, the car bleeps loudly to warn the driver. There may also be a visual display of how far away the obstacle is.

Modern cars have several cameras in the bodywork, looking far beyond what a driver can see from the driving seat directly or using mirrors. Sensors now use computer vision instead of radar, so can analyse what's around and even under the car and take more appropriate action than just bleeping. There are also other sensors involved in parking than just in rear bumpers. They now don't just detect vehicles and pedestrians directly behind a reversing car, but those in the distance that are likely to end up behind it during the manoeuvre.

Sensors can also spot vehicles in the blind spot over a driver's shoulder.[74] If a driver tries to change lanes while something's in the blind spot, the car won't just bleep loudly, it will also vibrate the steering wheel, and gently steer the car straight ahead to avoid a collision. The driver can over-ride this if they want, but if they've not noticed the approaching hazard, the car's action will prevent a collision.

The AI in the car also behaves the same way if a driver drifts out of their lane. In this case, the sensor used is an intelligent camera checking the car's position against the white lines on the road. Such cameras can monitor a driver's eyes for warnings of drowsiness, such as excessive blinking, alongside other signs such as the way they're steering. The car warns the driver if they seem sleepy, suggests a break, and can even direct them to a convenient place to stop.

There are several other examples of AI-enabled active safety features in modern cars, which we won't go into here. Examples include emergency braking, cruise control and self-parking, all of which can use combinations of sensors, computer vision, IoT device control and minor robotics to control the car based on changing circumstances.

Navigation, Entertainment and Communication

We've already covered AI navigation in some detail. What's worth highlighting about AI navigation and entertainment in cars is that much of the data they need is obtained from outside the car, via the Internet. Unlike most of the other examples of automotive AI, these examples need live updates, such as streamed music and traffic updates.

The most common way of achieving this connectivity is through a mobile phone, usually using Bluetooth or a wired USB connection. The other option is a SIM card slot in the car, the same kind as used in mobile phones. The advantage of a dedicated SIM card is not just convenience and cost, but also the possibility of using faster 5G mobile networks, which are being rolled out in many countries.

Once a car is equipped with a SIM card, it becomes a big mobile device in its own right, with more processing power than a typical laptop or mobile phone, and plenty of space for more sensors and other technology.

The Role of Human Drivers in Modern Cars

A modern car is a powerful AI ecosystem, filled with examples of ANI, receiving and generating huge amounts of data. It started life as a horseless carriage—a powered alternative to using animals to transport people and has evolved over a century or to do that very well. But if you take a fresh look at what it's capable of today, it has much more potential than simply doing a better job of what cars have traditionally done.

In particular, as you consider the combination of navigation, sensors and computer vision, the role of a driver becomes an interesting question.

Once we're capable of creating the driver aids and safety features, it's a relatively short step to extending that so the vehicle drives itself. Clearly there are very tough safety criteria to meet, and there'll be situations which probably only human intelligence can handle. And of course, there's the human factors involved in accepting such a seismic shift in what we're familiar with. Nevertheless, it's worth considering why someone may feel a human driver is a better or safer way of controlling a car that could technically drive itself.

Proponents of self-driving cars have persuasive evidence and arguments that the human driver is the weak link in cars today. They believe that committing to self-driving vehicles will create safer roads, faster journeys and less pollution. There are, of course, many counterarguments.

We still live in a world where autonomous vehicles are not allowed on most public roads. But increasingly, 'when' that will change seems a more realistic question than 'if' it will.

AI in Taxis

When people name areas where AI has made a major change, a common answer is Uber, the online taxi app. The business model for Uber, and competitors like Lyft and Ola, is a fleet of cars or other passenger vehicles (such as autos in India or tuk-tuks in Sri Lanka) in an area, waiting for customers to hail them. The customer hails the car using a mobile phone app, so the app knows where the customer is. It uses AI to find all the cars nearby and sends them a message with details of the customer's location and destination. The first car to accept the request is given the job, and the passenger is sent a message that the car is on the way. Once the car arrives, it takes the passenger to the destination using an AI-generated route, and after the passenger arrives, online payment is taken based on the journey time and duration.

The whole service is infused with AI. For example, it's used to estimate to the customer how long a car will take to reach them, and the duration and cost of the journey. AI is used by passengers to rate journeys, so other passengers can choose or reject different drivers if they want, based on feedback. Similarly, drivers can rate passengers.

Safety features[75] are becoming increasingly sophisticated. For example, a simple one is a button on the app so that passengers can share their progress and driver details with a friend. A more intelligent feature used by Uber in India involves computer audio in the Auto (not the customer's app) that constantly listens during the journey for words or noises that might indicate passenger's distress. If, say a scream or cry for help is heard, the AI monitoring the sounds in the car automatically places a call from the taxi customer service team to the passenger, to check if they are ok. They also call the driver, and if things don't seem right, they know the exact location of the vehicle.

In some markets, Ola has a couple of features so far unique to its service. For example, AI sends a numerical code to a passenger when they've accepted a ride, which needs to be entered into their phone when they get into the car. This is to make sure they're getting into the actual taxi despatched to them, and not an imposter. If you listen to music or watch a TV programme using the Ola entertainment system[76] in one cab, when you get into another it'll ask if you want to continue watching or listening where you left off at the end of your last journey.

These are all clever, convenient features, none by themselves ground-breaking, but collectively transformational compared to traditional taxis.

The other thing you may have noticed is that most of the references to taxis have been to 'cars', not 'drivers'. That's because the companies behind these services are all expecting to run fleets of autonomous vehicles eventually. So, they've designed everything to work with or without human drivers where possible, or with minimal change.

The question they're confronting us with is whether we'd have more faith in a human or AI taxi driver. If our loved ones were in a taxi home, which form of driver would we feel was safer? Not necessarily as simple an answer as it used to be.

AI in Trains and Planes

The AI in public transport systems, typically trains, aeroplanes, subway and buses, has probably had the least visible impact on the public, even though they already rely on it so much. That's because it's 'behind the scenes', so we only see its results, not its use.

Maintenance and Repair

One of the biggest costs in public transport systems is repair and maintenance of vehicles. For example, an airline can spend up to a fifth of its direct operating costs on this, nearly half just on engines.[77] It can cost between 1 and 10 million pounds to remove an aircraft engine and return it to the manufacturer for servicing. So, the kinds of AI technology used to improve car maintenance, such as IoT and computer vision, are much more valuable for large public transport vehicles.

AI is used extensively[78] to monitor and diagnose mechanical components, and optimize their operation. It's also used to simulate use over time, looking at different operating conditions and the effect on cost and safety. This simulation is particularly complex for aircraft.

Scheduling

Most people's view of public transport quality usually relates to service frequency and punctuality. Ensuring accurate, achievable schedules is a major contribution of AI to public transport.[79] The challenges vary by transport type, but most involve intelligent analytics in solutions.

For trains, the calculations and algorithms involved are complex, especially for large networks and busy routes. Updates are usually done in small steps, whether with or without AI support. Where major changes have been attempted, even the most sophisticated AI has yet to provide certainty that new routes will run smoothly from day one.

The AI involved in train scheduling[80] is based on a complex set of mathematical equations that represent the position of every train on a fixed network as the day progresses. Difficulties arise with unexpected changes like bad weather, breakdowns and other unplanned incidents. Just one train in the wrong place has a knock-on effect on the rest of its journeys for the day. It also affects other trains, such as faster trains that should have been ahead of it, but will now need to follow it, and so have their own schedule affected. There's also a problem for passengers relying on connections, as they may miss their planned changes. Finally, the train may be in the wrong place at the end of the day for the start of the next day's schedule.

This is just what happens to one train that's delayed. If there's a problem across an area, say a breakdown closing a piece of track or a security incident closing a station, the impact is many times greater.

The power of AI used in scheduling systems is all about representing the complexity of the network and solving the problem of how to keep it running as well as possible. Sometimes this involves difficult judgements such as cancelling a train to 'unblock' a route for many passengers, despite much greater inconvenience for a few.

NINE TO FIVE: AI IN BUSINESS

AI in work and business is difficult to explore briefly, because of the extreme diversity in how it's used. From robotics in factories to product recommendations in e-commerce, there's not an area of industry and commerce that hasn't been touched by AI and ML.

As we saw in the history of AI, this has been driven by the desire for investors and shareholders to make money from their AI spend, and the fear that others will make more money. For some sectors, rapid adoption and development of AI in business is no longer about innovation and differentiation, but a matter of survival.

There is plenty of readily available material that describes the details of its many applications. Rather than trying to add to those here, we'll instead take a higher-level view, to illustrate the breadth and diversity of how AI has transformed business.

We'll do this using a simplified, generic description of what most businesses do, based very loosely on Michael Porter's Value Chain model.[81]

Selling Things

All companies that make money do so by selling something. That may be a physical product (such as clothes or soap), an intangible product (such as an investment asset or insurance policy) or a service (such as household repairs or financial advice). AI is used throughout all versions of selling. Some aspects, such as advertising and marketing, use AI in similar ways across all three. Others vary tremendously according to what's being sold, and also whether the selling happens in physical stores or online.

Marketing and Advertising

AI is at the heart of influencing and persuading large groups of potential customers to prefer one product or brand over others.[82] Details vary by advertising 'channel', that is, the medium used to communicate with them, such as newspapers, TV or online. But the principles of AI involved are the same.

The problem being addressed is deciding which customers to target, what messages to send them to influence their future purchases, and how to evaluate the effectiveness of such efforts. In AI terms, this mostly means predictive analytics and forecasting. The data available includes customer profiles, customer behaviour, marketing activities and changes in product purchases, or at least customer interest in product purchases. For example, an online retailer will analyse shopping patterns for all its customers, perhaps concentrating on a particular subset such as affluent young men in London. Based on those patterns, AI may use predictive analytics to identify particular types of product they are more likely to buy in a given set of circumstances. Say, as the weather gets colder, those who buy a particular type of warm jacket also are more likely to buy a light raincoat. The company may try a marketing campaign to promote those raincoats to anyone looking at the warm jackets and offer an extra discount if they put the jacket into their shopping basket. Afterwards, the AI could use this data to not just evaluate the effectiveness of the campaign but improve its algorithms to identify other products affluent young men in London are likely to buy.

This is obviously quite a simple example, but AI allows much more complex patterns to be uncovered and used to design campaigns. Another extension of this is using AI to analyse patterns and make buying predictions for an individual rather than a group of customers.

There are many more examples and applications of AI in marketing and advertising, which we won't get into here. Some of the most interesting are in online advertising,[83] such as deciding where to place ads and what kind to use. Google and Facebook are two of the largest earners of online advertising revenue. It's no coincidence that they're among the most advanced users of AI in the world.

Physical Stores

When selling involves a customer visiting a store to buy products, AI is used extensively to make that process more effective and efficient.

Again, it's impossible to be comprehensive and brief without being superficial. Here are three areas worth highlighting.

The psychology of store layout and design is a complex art, with customers reacting strongly to changes to how items are grouped and even which shelf products are on. Supermarkets have used data about footfall, purchases and customer feedback to improve selling for years, but AI allows them to do this better.[84] Computer vision and other sensors can track the movement of customers around stores, and analyse how they choose to move between aisles or departments. By adding purchase data to this, retailers can then start to look at which layouts generate higher sales. Even simple things like queue lengths, availability of baskets versus trolleys and positioning of signage can affect margins. AI allows for incredibly detailed analysis of every conceivable factor in a store visit and helps make minor and major adjustments to improve sales.

In some types of store, finding the right product can be a problem. Examples include DIY and clothes. There's been plenty of work on using AI to help customers find what they're looking for online, and versions are now appearing in physical stores. Fashion retailer Rent It Bae has equipped its Delhi stores with voice-activated screens to allow customers to search for products.[85] US hardware store Lowe's introduced robotics in-store,[86] creating a robot that roams the aisles, asking customers in natural language what they need, and taking them to the right location.

Our final example is for more complex buying decisions, using AI to provide the guidance and advice a human assistant would previously have offered. Examples here include cosmetics, accessories and even clothes. Stores like Lush,[87] Sephora[88] and Lenskart[89] (in the UK, US and India, respectively) have experimented with in-store technology that uses computer vision to scan customers' faces and suggest makeup or glasses that suit them. In an even more advanced example, Japanese fashion store Uniqlo experimented with a kiosk that used AI to recommend clothes using brainwave detection! The in-store kiosks analysed a customer's response to seeing a product, using sentiment analysis to judge how they really felt about it.

Online Stores

E-commerce is probably the area of business where customers are most aware of the influence of AI. For most online shoppers, personalized recommendations are an expected part of the experience, and they're becoming increasingly familiar with variable pricing. People buying travel products online, whether airline tickets, hotel rooms or even taxi rides, expect the price to change depending on demand and other factors. Of course, AI is responsible for calculating those price changes today.

Many shoppers now understand the logic behind how AI sets discounts, so actively search for discount vouchers, or wait to see if they're offered price reductions after browsing. Some customers try to 'game' the AI by deliberately leaving the site after looking at a product they want, in the hope of being offered a discount when they return. In other words, AI has learned to emulate the human salesperson behaviour of chasing behind a customer as they leave the store, trying to persuade them to buy what they were looking at.

A growing challenge for online customers is finding the exact product they want amongst overwhelming product choices. AI uses increasingly sophisticated search algorithms to try to better predict what someone is really looking for. This involves a combination of better natural language recognition, data about the individual customer, and data about what other customers look for, choose and buy when they use search terms.

Another approach is to move away from description-based search completely and use computer vision to find products based on what they look like. Neiman Marcus[90] is one of many doing this, allowing customers to find a product by uploading photos of the kind of thing they're after. The online store then suggests products that look similar, based on colour, shape, style and other factors. Through ML, the quality of such algorithms improves with use. Other companies have extended visual search to recommend matching, complementary products. For example, if you look for a particular dress visually, the recommendation algorithm will use visual search to suggest matching handbags, scarves or shoes.

The final area of AI we'll cover in this section could fit into several others, but it fits here because most people experience it when trying to make an online purchase. And that's fraud detection.

Fraud is something that businesses have grappled with since they first existed, and modern payment methods make it easier than ever for criminals to attempt. The main forms of modern retail fraud, apart from physical theft, include making payment using stolen cards, getting products delivered to someone other than the customer, and exploiting the returns process to somehow make money. Behind the scenes, there's also fraud in the store or warehouse, whereby employees systematically interfere with orders for personal gain.

AI can help with all of these, but its main contribution to fighting fraud is in the detection of stolen cards.[91] The way this works is using the AI technique of anomaly detection, to spot exceptions to transaction patterns that indicate something untoward. This might be something obvious, like the same card being used in different places at the same time, or it might be its use for goods the real customer doesn't usually buy. Variations on the last of these include a sudden spike in high value purchases, or many frequent contactless purchases.

Processing Paperwork: RPA

Processing paperwork used to take up most of the time of most employees in many industries. From processing orders to managing complaints, many industries built complex businesses that ran on paper. Today, much of the information that used to be on paper is now held in computers. But the work is still based on computerized versions of the paper processes, and even the most computerized industries, such as banking, still spend a fortune managing paper.

Document management is a major area of AI application,[92] and most large businesses that handle lots of paper routinely use AI to manage it. Most of this is a combination of scanning and OCR to digitize the printed material, and intelligent document management to store and retrieve the digitized content. Many businesses, such as insurance, receive paper forms with content that needs to be understood and analysed, for example claim forms where the customer describes an insured incident such as a car accident, or a policy application in which they provide details that are used to calculate a premium.

AI extracts this information automatically, and you may be surprised by the amount of automated understanding and subsequent action that it can handle without human involvement. This is done by NLP.

When heavily paper-based businesses introduce computers, one of the most common goals is to reduce the amount of paper, entering as much information as possible directly into computer systems by people using keyboards. This includes customers as well as employees. While this reduces or removes the need for AI to process physical paper documents, it doesn't change the need for the other part of its use: automating the process that used to be paper-based. The use of AI in such processes is what we refer to as RPA.[93]

The reason RPA is considered an example of AI is because of the need to intelligently make decisions, rather than simply pushing information from one step to the next. There's a fine line between the two, and some vendors of simple process automation software take liberties in describing their product as intelligent.

In the early days of process automation, where little intelligence was involved, something like sending out an invoice could start an automatic process. The system would check payments received every day, find anything with a matching reference and amount, and change the status of the invoice to 'paid'. This would then automatically flag the transaction as ready to be posted to an appropriate ledger. For something as sensitive as financial transactions, there would often be human intervention involved to check the transaction and confirm the posting to the ledger. Computer automation meant they would do this with a single button-push or mouse click for a day's transactions, but it was still a person pressing the button to do the posting.

RPA uses AI to take this idea to its logical extreme, with the computer doing all the checking, posting and other processing required. When everything is working fine, an example like this is pretty straightforward, with little intelligence apparent. The AI in RPA kicks in more visibly to handle less straightforward processes and exceptions.

Hence, the R in RPA is describing the concept of software 'robots' that are intelligently performing the kind of activity that would have been done by people using paper many years ago, and at screens more recently. There are endless examples of RPA in use today, and most of us will have experienced an RPA process without realizing. Here are a few that illustrate it well.

Customer Order Processing

Once you click 'buy' on an online store, RPA is likely responsible for everything that happens between then and receiving the product, with the exception of a person leaving it at your home, and possibly a person packing it at the warehouse.[94] The software robots look after sending the order details to the warehouse, updating financial and inventory records, and organizing the shipping.

Replying to Customer Emails

Most emails received by customer service departments are asking the same kinds of questions, typically asking for updates on an order, querying an order detail such as price or shipping, or making a general enquiry. Some will be exceptions to this, and some will require a complex answer. But most involve retrieving a customer or product record from a system and replying with details or sending standard information that answers a standard question.

The way RPA works here is to use NLP to automatically read every incoming email and pick out the ones that fit into any of the categories it's been designed to handle. The rest are passed to humans. The ones about orders or customer accounts are analysed for relevant account details, and the appropriate record is accessed. It then uses NLP to compose a natural language answer and sends it back to the customer in an email. The more general queries are answered by inserting the relevant text from libraries of standard answers and sending that back as an email.[95]

In both cases, the AI functionality here is similar to the chatbot example we've seen before but used to create an email rather than a chat message.

Payroll

Large companies used to have whole departments of people whose primary job was to make sure employees got paid the right amount on the right day of the month. As well as regular salary payments, and these teams would look after calculation of tax, other deductions, bonuses and expense reimbursements. Whenever there were legislative changes, such as new income tax rates, their job would be to make adjustments to payments. And of course, these teams would ensure payslips were printed and sent out, and the correct monies paid to employees.

Unsurprisingly, AI and RPA specifically means that such departments are smaller or even no longer required. AI can do the same job faster, more accurately, and more cost-effectively.[96]

The rules and calculations involved may vary by employee, and data will be required from other systems. For some businesses, there may be a need to calculate hours worked from hand-written timesheets. Expense calculations can be fiddly, and receipts need to be digitized and stored with other tax records. It's not a simple exercise, but in AI terms is nothing more than a mix of intelligent application of rules, perhaps with some document management and handwriting recognition.

Insurance Premium Quotations and Renewals

We've seen that RPA can be used to process new orders for a product or service. The exercise is relatively straightforward for a simple e-commerce purchase but requires more sophisticated intelligence for insurance premiums.[97]

The first part of the process is obtaining all the information required to give a quote. This is usually done online, or perhaps via a hand-completed form received in the post. Some of it forms simple inputs to the quotation calculation, such as age, postcode, number of years since any previous claim and so on. But much of it involves risk calculations, for example, the probability of a car accident, burglary or illness, depending on the kind of insurance. The risk assessment is another area for AI, something akin to weather forecasting in how AI does it. But the actual calculation of risk and associated premium is beyond RPA and is an input RPA receives.

Once the RPA for premium quotation receives the price, it can complete the rest of the process itself, preparing a letter with all the right customer details, description of cover, terms and conditions and so on, and send it to the customer by email or post. If the customer accepts the quotation, it's converted to a completed order by RPA, which sends out the paid-up policy details.

There are endless ways RPA can be used in almost any business, certainly large ones, but even small ones. What these examples have hopefully shown is that the AI involved in RPA isn't usually very dramatic or even advanced, but it's typical of the way AI is being embedded into life all around us.

Delivering Things

Logistics companies like DHL, FedEx and UPS have always relied heavily on technology and are renowned for their expertise in subjects like navigation systems and demand forecasting. AI is now a fundamental part of these.[98] However, in recent years their supremacy has been challenged by newcomers to the business of delivery to customers, who have used AI to drive this, particularly analytics, forecasting and robotics.

At one end, large retailers have become delivery experts themselves, because innovation in delivery has been a differentiator in their core business. Features like same-day delivery haven't been possible if a retailer is outsourcing delivery to a third party. Amazon is recognized as a leader in this, but other e-commerce giants such as Flipkart are also exceptional logistics firms in their own right. In the UK, the first retailer to offer same day delivery was actually Argos,[99] not Amazon.

Meanwhile, other companies with relevant skills have entered the delivery sector from a different direction, especially for local deliveries. Firms like Uber have taken their expertise in managing fleets of local vehicles to get to customers quickly, and applied it to delivery, starting with takeaway meals.

Making Things

We don't have space to look in any detail at how AI has reinvented the modern factory, but anyone familiar with them will know that factories have been highly automated for years. As long ago as the

1990s, cars were put together by computer-controlled machines that aligned metal panels and welded them together. So-called hi-tech manufacturing, such as electronic components, involves factories that look more like laboratories than industrial buildings.

AI has been used to control the machines that make the goods for a while, using IoT and computer vision. Its presence is now even more visible in the robots that move items around the factory, from machine to machine. To all intents and purposes, these are autonomous vehicles, but designed to carry products and components rather than people.

This technology is even more important in warehouses, where stock is moved from storage unit to distribution areas and packed without any human involvement. AI doesn't just control the machines and vehicles. Through computer vision and intelligent automation, it also keeps track of where things are, and calculates the optimal way of grouping things together for delivery efficiency. Boxes and items that used to be identified using barcodes are now identified by computer vision as well. AI also uses IoT sensors for things like weight and size of items to calculate the best way to pack them, and spot mis-located items.

Planning, Analysing and Managing Things

A critical activity in all businesses is planning and managing activities, usually after analysing data such as sales, productivity, costs and other information that affects business performance. AI has turned this type of work on its head and is now taken for granted by most businesses.

The origin of this technology is the spreadsheet, originally a tool that automated basic arithmetic, to prepare financial information more quickly than calculators. The spreadsheet is also most people's first and perhaps only experience of handling business data, and is today a sophisticated tool for manipulating, analysing and presenting data. AI takes this to another level, including its use within spreadsheets.

Spreadsheet users may be familiar with its inbuilt formulae, such as summation and averages. They might know that the spreadsheet automatically identifies sets of numbers that it thinks should be added together, or series of data that can be extended. For example, if you label three columns Jan, Feb and Mar, it will automatically label subsequent columns Apr to Dec and beyond. When they first appeared, these were impressive features, but are now trivial. What they show is the first common appearance of AI in analysis, and why it's easy to not even realize it's present.

The type of intelligence involved is problem-solving and reasoning, and the types of AI used include predictive analytics, forecasting and even NLP (if the data to be analysed is provided in natural language). The simple principles that first appeared in spreadsheets have been extended to every type of analysis a business does and enable AI to solve incredibly sophisticated problems.

The most common application is probably in forecasting sales and financial figures and identifying what's happening in areas that could be improved. By creating models in mathematical language of what happens in a business, it's possible to simulate the impact of all sorts of scenarios. Typical examples of these are advertising campaigns, pricing changes and opening new stores. The sophistication means the AI can take account of factors that were previously too complex to consider, for example, the impact of traffic changes caused by a new store on footfall.[100] Or perhaps the possible psychological response to a new price across all customer segments.

The power of AI analytics in planning doesn't just come from the algorithms. It's also because of the ability to take account of huge volumes of data, including relevant publicly available data that isn't part of the business. For example, weather affects store visitor numbers and social media trends can impact product popularity.

There are endless ways AI is and can be used to improve the analysis, planning and managing of activities in business. To see examples and find out how they work, there's much material available elsewhere on the subject. Here, we've only seen a glimpse of what's possible and what's behind it.

Employing and Managing People

The final area of business we'll take a quick look at from an AI perspective is HR (human resources), what used to be called Personnel. It's the function which brings new talent into organizations, looks after their various needs while working for the business, and manages their exit. It's also the function which develops skills of individuals and helps ensure that behaviours are within acceptable standards.

At one level, using AI to manage human resources is comparable to other types of management and planning. There's data about the asset (in this case people), there's data about performance, and there's data about expectations and plans. So, AI can be used in recruitment, appraisals and the planning of training and incentives. From a computer perspective, the fact that the subject is a human being only really affects the type of relevant data, and the rules used to design ways to improve the performance.

Of course, the challenge is that people aren't just assets, and we have to be much more careful about what we allow the AI to do. While it might be ok to simply replace machinery when its performance degrades or a better version is available, we're not allowed to do that with an employee. There's also the fact that employees respond to how they're being treated, and so models to predict behaviour are much more complex than more typical AI business problems.

The use of AI in HR is an ethical and legal minefield. As employers are able to use AI in increasingly innovative ways to gain data about their staff, the importance of rules on how to use them also increases. The root difficulty is two-fold. First, there are very valid reasons for capturing certain data, but once it's available, it could just as well be used for other purposes. Second, some of the ways firms try to obtain this data are intrusive and not always transparent.

Employees now typically carry identity badges for security, often using them to gain access to buildings, internal doors and even resources such as computers. This means data is available on how they spend their workday and, in theory, could be used to track their whereabouts to the minute. AI security software ensures employees are not doing anything untoward with company data. But it's a short step to extending this tracking to their keyboard strokes, and the contents of their typing. AI computer vision in security cameras analyses what a person is doing to spot suspicious behaviour for safety reasons. But in doing so, it's recognizing activities like talking to others, reading and answering the phone. In the same way, it's now accepted that work phone calls may be recorded, especially to customers, but this means data is available for NLP to analyse the contents of employees' phone conversations.

With all this data available, and justified on security and other grounds, there's no technical reason why it couldn't be used by AI for more debatable purposes. As a business, this could be useful to improve the performance of staff. But as an employer, there are other obligations and constraints. As we'll see later, what's not clear is who is responsible for deciding where the lines lie in such situations, and what happens if they get crossed.

This is not theoretical. Like the rest of this chapter, the examples are already starting to appear, or are coming soon. In early 2020, Barclays Bank introduced a new employee monitoring initiative in which AI tracked data about how much time staff spent at their desks. There was an outcry, and the initiative stopped.[101] But at the time, the bank didn't clarify if it had stopped or merely paused it. Nor was it apparent whether they thought there was anything wrong with the initiative in principle, or if the issue was its acceptance.

These examples illustrate how AI is starting to change many accepted ways of approaching normal day-to-day activities. This has many benefits, but also opens up dilemmas, challenges and broader questions. That leads us to the remaining chapters of the book, where we explore the questions AI and ML pose about our future.

5 THE RISKS, CONSEQUENCES AND DILEMMAS AI BRINGS

'Do go on', he said. 'There's nothing I enjoy more than listening to a highly trained intelligence leapfrogging common sense and coming to the wrong conclusions. It gives me renewed faith in parliamentary democracy'.

—Tom Sharpe (*Wilt on High*), Author and Satirical Novelist

Any new technology brings with it challenges and dilemmas, some significant, others merely inconvenient. To date, we've usually found answers to the implications. Some, such as electricity, have reshaped society. Others, like e-commerce, have also had major impact, but are only better ways of doing previously familiar activities. AI does both.

To date, we've only really seen the upside of AI and are yet to experience many problems. Perhaps more accurately, we've started to see the first negative consequences, but haven't yet fully appreciated their nature and implications. Public concerns about AI remain relatively theoretical and tend to focus on the longer-term future.

This chapter aims to adjust any perceptions you might have that the difficulties we'll face because of AI are a long way off. It describes what these difficulties might be and how they arise.

WHO DECIDES THE RULES?

There are risks and issues with almost every example of AI covered so far, but they're mostly typical of the kind of problem that arises with all new technology. It's not that they're unimportant, but history shows us that this kind of issues gets resolved, often by those developing it. If they didn't, it wouldn't survive. More importantly nowadays, it wouldn't make money.

However, AI also brings with it a series of larger issues that are beyond the purview of any individual application, researcher or even institution. The roots of many of these lie in intractable questions, like what 'control' means for an intelligent machine created by human designers, especially when it can learn, and one day may even possess a mind.

What Does 'Control' Mean in AI?

Our ancestors who first conceived of AI imagined thinking machines to be slaves and servants, in control of a natural expectation of superior human masters. Today, we consider AI to be more tool than just slave. But because it has intelligence, the question of control is still important.

Control in an AI context means three things:

- Designing how AI comes up with answers and results, and therefore controlling that it works correctly.
- Deciding how the results of AI get used in practice, and so controlling the use, misuse and abuse of what AI achieves.
- Protecting those who might be inconvenienced or even harmed by AI; hence, controlling the consequences of AI when something goes wrong.

Each is a different type of control, with different approaches to achieving it. The biggest difficulty with all of them is the lack of any agreed ‘rules’, and associated authorities to enforce them. It wouldn’t be impossible to specify the relevant issues and formulate suggestions for managing them. The challenge of course is agreement. Hence, the emphasis of this section is not on what the rules should be, but who decides them.

Similarly, deciding where the authority lies to enforce any such rules is even more complex. It brings into question the ability for laws to work on problems that would have been inconceivable when they were written. Relying on businesses and regulators to enforce rules is also unlikely to be straightforward for many reasons, not least their motivation.

When it comes to controlling the use of AI, it’s hard for those outside an organization to know how AI is being used, and therefore if it’s being misused. For example, if less affluent customers are frequently declined loans, it’s not easy for them to find out if it’s because their credit-worthiness has been fairly assessed, or if factors such as race were used.

What makes it even more complex is that those making the decision may not even know the factors considered. Through oversight or human error, AI might not have been prevented from considering socioeconomic or personal data inappropriately. If it’s allowed to look for patterns in such data, it may use them to ‘improve the performance’ of the loan decisions by assessing loans in a way that shouldn’t be permissible. Even if controls were in place during initial design to prevent this, ML means that it doesn’t automatically follow that those controls would continue to work if the calculations change over time.

Finally, looking briefly at controlling the consequences of AI problems, what we’re really discussing here is enforcement and remedies when (as yet non-existent) rules are breached. One option is to adapt existing laws, regulations and rights to deal with AI. The other is to create new ones designed for the unique challenges of

AI. Given how long it's taken to deal with new challenges such as data privacy regulation, it's not clear how we'll find appropriate, internationally acceptable answer to AI's challenges.

A Human Analogy: Medical Care

We can try to bring these ideas to life with a non-AI equivalent, and then see how the same things would work with AI. An area where these concepts apply directly is medical care. The equivalent to issues about control include: how we know and ensure human doctors come up with the right results; whether we can be confident they're doing the right thing with those results; and what happens if they get things wrong.

We rely on human doctors because all of these issues have been thought about and addressed for us, and we have confidence in those that looked after those issues. For example, we assume and can check that a doctor is appropriately trained and has been independently certified as competent. We also know that they use instruments and tools (e.g. blood pressure monitors, labs to analyse blood) that are reliable, and have also been certified for use. If, despite this, things go wrong, the doctors, instrument-makers and medical institutions are regulated by bodies that we broadly trust.

Most of us don't check any of this when we visit a familiar hospital, but we may well do the first time we see a new doctor. We might not check their competence, but if we wanted to, we can check their qualifications, professional history and even feedback from other patients. Similarly, we don't usually check if their instruments or labs are reliable, but if we have concerns, we can check kitemark, regulatory approvals and other certifications.

Applying the Medical Care Analogy to AI

To compare this with AI, we now need a hypothetical discussion, because AI healthcare is still primarily only used to support human doctors. Our example is about an AI doctor that doesn't yet exist for patients to see instead of a human doctor.

Let's suppose you consulted such an AI family doctor, a talking machine trained in general medicine, designed to discuss and assess your symptoms, prescribe treatment for some ailments, or refer you to a specialist for others. You'd want to know the same three things you'd want in a human doctor, before even considering whether you'd allow AI to deal with your health problem: (a) Are they competent? (b) Are their instruments and tools reliable? (c) Are they effectively regulated?

And herein lies one of the biggest challenges for using AI in everyday life: Who should ensure the answers to those questions will be acceptable, and who should hold those people to account. These are critical questions for something as important as healthcare, and most people would care very much about them before switching from a human doctor. But we may not pay as much attention when the example is less stark.

To find a better answer, perhaps we should think about a slightly different question here, given our basic understanding of how AI works: not who ensures our questions are answered, but how they get those answers.

Let's go back to the hypothetical AI doctor which asks you about your symptoms, and decides what treatment you need, if any. To understand the AI involved, we'd break down the activity into smaller steps, find the intelligent ones, and look at the data, algorithms and ML involved to get and improve the best answers.

If we did that exercise, we'd find some NLP to understand our symptoms, and some form of intelligent search and prediction to assess what it might mean, and what to do about it. For us to be confident in the AI doctor, we'd want reassurance that the NLP designer did a good job of extracting medically useful information from our words. We'd also want to be very sure that the AI Search for matching our symptoms had access to a reliable set of symptoms, and the AI Prediction would accurately spot the most likely ailment. But we'd also want to know it hadn't overlooked less likely but more serious illnesses.

Obviously there's much more we'd need to be reassured about. So, the question we're really asking is: Did whoever approve this piece of AI for medical use know enough about both AI and healthcare to ask and answer the right questions on our behalf? As patients, we should also know how confident we should be in that.

This brings us back to the theme of this section: Most of us are unlikely to be able to check whether the AI we experience is reliable, safe and correct in what it does. What we may have to do instead is check the people who decide that it is and insist on ways to hold them to account.

DECIDING WHAT 'GOOD' AND 'BAD' AI MEANS

There are many dilemmas, risks and challenges associated with AI, and the list will only get longer as we find new ways of using it. One of the most difficult is an age-old philosophical conundrum about the meaning of 'good' and 'bad'.

'Good' and 'Bad' in Philosophical Terms

There's a long discussion to be had on what happens if AI gets into the 'wrong' hands, people like fraudsters, thieves and terrorists. We're not going to discuss that type of 'good' and 'bad' AI, because that's more about security than AI. We're going to assume that all humans involved are well-intentioned, and there's no malice or misuse to worry about. That restricts the discussion to a slightly more manageable problem: How AI deals with difficult choices, where 'right' and 'wrong' aren't necessarily obvious?

An old example often quoted by philosophers to illustrate this issue is the Trolley Problem,[102] a thought experiment about a choice between inaction and action, both causing people to die. If you're not familiar with it, imagine standing by train tracks, along which a runaway trolley is speeding. In the distance are five people who'll be killed when the trolley hits them. There's a lever near you which will change the trolley onto another track, on which one person stands. If you do nothing, five people will die through no action of yours; if you pull the lever, only one will die, but as a result of something you actively did.

AI doesn't offer any answers to the underlying philosophical dilemma, but the new problem it raises is how you represent and evaluate the choices. We've seen that a key element of AI is using mathematical language to represent activities, data and decisions. So AI relies on maths, statistics and logic to make such decisions.

It's probably clear how a data scientist could represent this dilemma as numbers of lives lost. Simply put, one life lost is 'better' than five, so the maths justifying pulling the lever is clear. Designed this way, AI would pull the lever because the answer is mathematically obvious.

However, the philosophical issue is about taking life by actively doing something versus allowing life to be lost by inaction. There's a school of thought that says that actively taking life could be considered 'worse'. Rather than debating this, the point is that in AI, both factors can be represented mathematically: How many lives are lost and also how 'active' the decision to take life is.

If the representation includes both factors, the data scientist has to decide how big each factor is relative to the other. For example, they could make an active decision to take life 10 times 'worse' than a passive one. If they continued to represent the other factor as the number of lives lost, then there's a simple piece of arithmetic now involved: Actively causing the death of one person is twice as bad as passively letting five die, so the lever should not be pulled.

If the arithmetic isn't clear, the active choice, pulling the lever, is 10 for the active nature of the choice times one for the number of lives lost, giving a score of 10 ($10 \times 1 = 10$). The alternative score is 1 for the passive nature of the choice times 5 for the number of lives lost, giving a score of 5 ($5 \times 1 = 5$), half the other score.

Which one is right or wrong is not the point of this illustration. That was the purpose of the original Trolley experiment debate, and of course there's no right or wrong answer. Our purpose is to show why it's important to understand what is being represented by maths behind a piece of AI. As important is who's making what judgements and decisions when designing the AI that will apply their logic in practice.

AI is already affecting so much of our lives. Even if we assume its designers are doing what they genuinely believe is right, the point is that some of these decisions shouldn't be theirs to make. But unless we realize what's involved, we may never even think to ask about such matters.

Obviously, the example is an extreme one, but the principle applies in all applications of AI. For example, if AI is used to make a decision on whether to pay out an insurance claim or allocate a place in a class, it's still important to know that 'right' and 'wrong' answers are being decided using logic that's been designed appropriately and acceptably.

'Good' and 'Bad' in Business Performance Terms

Now that we've looked at extreme examples of 'good' and 'bad' in AI, we can return to more the more prosaic world of AI in business, where simply trying to perform a business activity better can create dilemmas. Of course, these won't (or shouldn't often) involve risk to life, but the AI challenges can still be complex. This is because AI is a powerful way to improve the performance of an activity. If there isn't clarity on what ‘improved performance’ means, we can quickly create difficulties that are hard to spot and can be harder to undo.

Deciding what ‘improved performance’ means is central in determining the scope and purpose of any piece of narrow AI, and the basis for the logic that human designers build and algorithms they select. As they translate ‘improved performance’ into AI goals, measures and designs, they are setting up how the AI will evaluate choices.

In AI business applications, a ‘good’ choice is one which will help improve the performance of the business activity. What this actually means depends on the definition of ‘improved performance’ given to the AI designers. Furthermore, the relationship between this and any individual step in the activity, such as acceptable risk factors to use when calculating an individual loan acceptance score, is invariably complex.

So, in AI business applications, ‘right’ and ‘wrong’ are nothing to do with morality or philosophy, or whether AI designers have ‘good’ or ‘bad’ intentions. It’s something much more grey; it’s a score that people designing AI have attributed to a factor, usually one of many, and generally part of a complex landscape of business objectives, priorities and pressures. And above all of these are overarching goals of the business that’s paying for the AI such as share price or sales figures.

This is why AI used in business has the potential to create unexpected and possibly unacceptable results, even though its designers may have sincerely used best efforts to solve a relatively neutral business problem. The very act of setting performance improvement objectives for AI work can build in ‘good’ and ‘bad’ influences on the results. As these are translated into details of an algorithm, then refined by ML, the link between those influences and the end result can become difficult or even impossible to spot.

The purpose of this section is to bring to life why the issue of 'good' and 'bad' in AI work is so important, as is the question of who decides what that means. This is especially the case when those decision-makers face complex combinations of goals, priorities and pressures, be they personal, commercial or even electoral.

IS IT TOO LATE TO ASK WHO SHOULD 'CONTROL' AI?

We've looked at the need for rules around controlling AI, and why that's both important and difficult. We've also explored how the concept of 'good' and 'bad' is not just a moral or philosophical question for AI. And we've seen how that leads to another aspect of control, deciding what measures are used to assess the results AI produces, and the implications for AI design, AI designers and those who commission AI work.

A thread through this chapter has been who should set such rules and controls, and how they could be made accountable for their responsibilities. Implicit in such questions is that there's a choice; this final section explores that choice, and whether it still exists in practice. There's an argument that the answers are already being formed, and the people shaping them have already selected themselves.

We'll do this by going through some of the possible sources of AI control and rules. These are the places we could potentially turn to in future to make sure AI stays within acceptable limits. This would all first need some sort of consensus on the issues involved, and their relative importance. This doesn't even mean answers to difficult questions, but agreement on what the questions are, and what should be considered to formulate answers.

Another part of the challenge of implementing controls around AI is the dynamic relationship among governments, citizens, experts, corporations and lawmakers. These are all stakeholders in the world that AI is creating, and they're all busy dealing with many other problems. With a few exceptions, it seems unrealistic to expect any of them to act alone and hope to get agreement from the others. As we've seen with nuclear power, climate change and freedom of speech, the more controversial, complex and important the subject, the harder it is to get broad alignment. And meaningful action is even more difficult to achieve.

Let's start with what we might hope would be the most powerful place for rules and controls on important matters: the Law.

What Can Laws Do About AI?

For laws to work, they need a framework of situations and regulations that can be tested in courts. There are existing laws that could apply to the potential problems created by AI, but their original context was far removed from a world of thinking machines. It's hard to see how most of them would be of practical use.

The other difficulty with existing laws is that they should include remedies: If someone has been wronged by an unlawful act, the relevant law should prescribe how to put things right, or make up for the damage caused. But the consequences of AI problems are complex, and many of the laws that might apply to AI are limited to financial remedies. Crudely speaking, with the amounts of money at stake for those involved in AI, these remedies may be a price worth paying.

History shows that laws made in haste are often poor and can even worsen a problem they're designed to help. So, the lawmakers are usually conservative for good reasons, and are long term and considered in their thinking. That doesn't make for speedy change, in contrast to AI which is now moving rapidly. In the short term, it's not obvious how the Law could keep up with AI and its potential problems. It's equally unclear how the Law might provide practical help resolving the issues of rules and controls for AI.

Can Regulators Control AI Effectively?

In some areas where the Law has been unable or unwilling to step in, regulators have attempted to provide controls, usually on behalf of consumers. Sometimes these are formed voluntarily by businesses in the affected industry, other times they've been set up by governments, backed by powers enshrined in law.

For obvious reasons, businesses prefer regulatory authorities to judicial ones, especially if they can be persuaded to allow self-regulation. Sometimes this works, other times it's a step taken unwillingly towards more stringent industry control. This is certainly an approach that can move more quickly than legal solutions but has its own challenges.

One difficulty is that regulators are usually formed for a specific industry, but the problems we're considering apply to any industry that uses AI. However, that's not insurmountable, and a cross-industry AI space could probably be defined in some form. There are also precedents for areas that cut across industries, such as data privacy. However, what that example also showed is that regulation needs 'teeth' to be effective, as early attempts to introduce data privacy codes of conduct didn't stop the worst culprits. However, rules and principles can be established through industry bodies, and create templates for subsequent laws if required.

Perhaps the biggest challenge to a regulatory approach to AI is that there's often no easy way to force businesses to let themselves be regulated. In some cases, there's already some legal compulsion in place, such as licenses to trade (as with banks and energy providers). But in the absence of this, it only takes one major player to refuse to sign up to regulation to undermine efforts.

The other significant difficulty with regulatory controls is that they tend to work at national or regional levels, and few regulatory frameworks apply internationally. It matters because the battle for AI supremacy has started to split along national and regional lines. In particular, the USA and China are leading global AI work on virtually all fronts. It's unlikely that companies in either country would agree to regulations without participation from the other. It's also questionable whether those governments would want their companies to agree.

Regulatory bodies may well start to appear and play a bigger role in AI work in some countries and on some subjects. But it's not clear what the path is from that to comprehensive industry regulation. And patchy regulation could turn out to be worse than no regulation.

Can Public Pressure Make a Difference?

In some areas, neither laws nor regulators have done enough to satisfy public opinion. The resulting public backlash has sometimes raised awareness of an issue, and in some cases it's forced change. However, on other subjects, public pressure has made no practical difference. Areas where the public voice has been loudest include nuclear disarmament, climate change and equal rights. The actual difference this pressure has made can be debated, but at the very least, it's caused governments and corporations to be more considered in their actions.

What these and similar topics have had in common is large groups of people who have cared enough about an issue to act. A second common feature is a relatively simple ask or action, generally stopping something. The ask may be unrealistic or aspirational, at least initially, but it creates a rallying cry around which people unite.

Given the benefits AI provides and how intrinsic it's become to daily life, it seems unlikely that enough people will care enough, for AI to become an issue for public protest. Another reason public outcry is unlikely is the relatively low public awareness of the risks, issues and consequences at stake.

Of course, those may both change at some point, but no one seems to expect it anytime soon. There's currently no well-known AI equivalent to 'Ban the Bomb' or 'Save the Planet', although a pressure group called 'The Campaign to Stop Killer Robots' has been building momentum since 2012,[103] and now has a wide international membership.

Do Governments Want to Control AI?

The short answer to whether Governments want to control AI is a resounding yes, and some don't hide that desire. Even those who espouse traditional forms of democracy don't make it clear how they intend to use AI. For example, much AI development was and is directed by governments towards defence and security.

AI in modern warfare is a terrifying prospect, but the picture is not as simple as AI being a military or intelligence tool. Economic power has become as important as military power, and AI is seen as a crucial part of future economic growth. No government, especially among leading nations, can afford to risk falling behind in AI capability, and all believe that AI could provide huge economic benefits. This doesn't just apply to countries with the largest economies or populations. Smaller countries have previously used technology advantage to punch above their weight economically.

So given the potential geopolitical importance of AI, it doesn't seem to be in any government's obvious interest to push for a level AI playing field. Nor are their priorities and goals straightforward to fathom.

Who's Left?

If you discount the groups we've looked at so far, there remains one place where rules and controls of AI may emerge, and have already started to appear. That's the large, global corporations leading the world's AI work. These consist of nine companies[104] from the USA and China: Alibaba, Amazon, Apple, Baidu, Facebook, Google, IBM, Microsoft and Tencent.

Known by some as the Big Nine, they are responsible for the vast majority of AI innovation on the planet, and control or influence most of the world's AI research. It's already been argued by several experts that these nine organizations have the most control over the future of AI, and its effect on the human race.

It's very hard to argue with that.

This chapter completes our journey of trying to understand what AI and ML are, and how they work. With this, you have enough for an informed, balanced view of this simultaneously fascinating, impressive and daunting subject, and what it may hold for you. It also allows you to speculate on what a future living with AI might look like, which is the subject of the next and final chapter.

6 END OF THE BEGINNING OR BEGINNING OF THE END?

By far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.

—Eliezer Yudkowsky, Co-founder,
Machine Intelligence Research Institute

Many experts have spent their careers thinking about AI. They've built knowledge about the past, present and future, and earned the right to be considered authorities on the subject. Here's what some of the more well-known of them have to say on what a world inhabited by thinking, learning machines might look like.

In distilling their perspectives into a few paragraphs, I've only covered their ideas relevant to AI, and the main messages and positions. But these are my interpretation, and the best way to fully understand what they've said is to read or listen to their own words.

Some of the ideas are complex, many are controversial. But all of them are valid and should be considered thoughtfully. The selection includes some of my personal favourites, but as it's limited by space, I've omitted many equally interesting and insightful experts. Those I've included cover varied backgrounds, including academia, business and journalism. We start with a couple of giants from the world of fiction.

THE FUTURE ACCORDING TO SCIENCE FICTION

Arthur C. Clarke

Sir Arthur Charles Clarke, along with Robert Heinlein and Isaac Asimov, was one of the 'Big Three' science fiction writers of the 20th century,[105] and winner of many awards. His fascination with space travel, combined with his physics education, meant many of his predictions, about space travel in particular, proved to be scientifically valid. In fact, he wrote several non-fiction pieces on space and rocket technology. One was influential in the development of today's GPS technology.[106]

He is considered highly influential in popularizing science and technology, and received a UNESCO award, the Kalinga Prize for this.[107] He was often asked to comment on space technology, including as a TV commentator for the Apollo 11 moon landing.

He wrote about space technology and its use but was equally interested in technology's impact on society. His books included descriptions of fictional technology that became reality, and he consolidated these into a series of essays, books and interviews. As well as GPS and satellite TV, he also described versions of the World Wide Web, online banking and smartphones.

The reason for his inclusion here is that machine intelligence was a constant theme through his work, usually described in terms of its application rather than its technology. This was most vividly brought to life in the movie *2001: A Space Odyssey*, based on his short story *The Sentinel*.

Described as one of the most influential films of all time,[108] one of the film's main 'characters' is a sentient computer named HAL 9000. HAL became one of the best-known examples of general AI, and arguably Superhuman AI. It, and all the technology in the movie, was inspired by AI scientists of the time. In fact, Marvin Minsky, one of the most respected people in the field then and since, was an advisor to the film. IBM was also consulted during its making, resulting in one of the first examples of technology product placement.

HAL is an example of what general or superhuman AI might look like and includes many forms of ANI that we've since achieved. These include NLP, facial recognition, playing chess, autonomous piloting, sentiment analysis, lip reading and even art appreciation.

What sets 2001 apart from much other work is its depiction of what happens when AI goes wrong. It brought to life the dilemma of a general AI that's been incorrectly designed. In this case, a programming conflict leads to HAL concluding that killing the human astronauts is the logical and correct action to achieve its goals.

The reasons it came to that decision are a vivid culmination of many of the issues raised in the earlier discussion of AI ethics. Essentially a pessimistic view, it shows one plausible route to AI turning on its human creators. However, it's only one view, and contrasts with the suggestions of Clarke's great rival, Isaac Asimov.

Isaac Asimov

Isaac Asimov is another of the 'Big 3' science fiction writers. Unlike Clarke and Heinlein, Asimov's passions lay closer to Earth, and his writing on robotics remains hugely influential. Anyone working in AI, especially in the 20th century, is likely to know of his work.

A biochemist, Asimov wrote on many topics, both fiction and non-fiction. His writing about AI was primarily around robots rather than computers. He personally knew Marvin Minsky, and was also a close friend of Gene Roddenberry, creator of Star Trek. In fact, he was a science advisor on the first Star Trek movie.[109]

His most visible contribution to AI is not as well-known as his other achievements: He is the first person to have used the word 'robotics' in print (in 1941). He later claimed he hadn't realized it wasn't a pre-existing word.[110]

His most well-known contribution to AI is his '3 Laws of Robotics'. These are an ethical framework for governing the actions of AI.[111] They are theoretical and fictional, but nonetheless a lucid approach to AI governance, not just applicable to robots, but to AGI. They postulate three prioritized rules, built into the artificial brain of every robot, that govern all their activities. The first, and highest priority law is that robots may not harm humans, including through inaction.

A commonly discussed concern about General and Superhuman AI is the risk that they turn against their human creators, perhaps for their own survival. Asimov's laws address this by including robot self-preservation as the third law, but only if it's not in conflict with the other two. The second law ensures robots always obey human instructions, unless in conflict with the first law.

As well as his huge influence on 20th century AI researchers, Asimov has been included here because he gave us an inherently optimistic picture of how AI could develop and stay benign. His work, whether credited or indirectly influential, is guaranteed to be part of how we reach future decisions on controlling and managing AI.

THE FUTURE ACCORDING TO SCIENTISTS AND ACADEMICS

Ray Kurzweil

Ray Kurzweil is known by many as a writer and futurist. What may not be so well-known is that his writing is built on a long and distinguished scientific AI career, including working with Marvin Minsky at MIT. A winner of the National Medal of Technology and Innovation, the highest US award for technology, he was named as one of the '16 revolutionaries who made America'.[112]

Kurzweil has written extensively on the convergence of human and technical capabilities, particularly the concept of the Singularity, described below. His outlook is positive, and his predictions about AI fall into two broad categories:

- How intelligent technology will become ubiquitous, pervading and improving every aspect of human life.
- How intelligent technology, including robotics, nanotechnology and biotechnology, will extend human life.

The Singularity

The Singularity describes the point in time when AI and other advanced technologies start growing at a rate that is irreversible, and no longer under human control. This results in dramatic changes to human life that can't be foreseen by simply extrapolating from using past rates of development.

Clearly, if this were to happen, the implications would be profound, and opinions on these tend to be polarized. Some see this as triggering the demise of the human race. Others believe it would create the possibility of a near-utopian existence. Kurzweil looks forward to what the Singularity would bring, and has proposed many examples of how life would improve on the way to the Singularity and afterwards.[113] He didn't invent the concept, but has probably done more than anyone else to make the public aware of it.

The Synthetic Neocortex and AGI

One of Kurzweil's most eye-catching suggestions is the creation of an artificial version of the cerebral cortex (or neocortex), the part of the brain responsible for higher order functions including perception, cognition, motor commands, spatial reasoning and language. In other words, much of the list making up intelligent activity.

A major difference between lower and higher intelligence in species is the relative size of the neocortex. The human neocortex is composed of six layers, and contains around 20 billion neurons, and around twice as many cells as a chimpanzee's. Kurzweil questions what would happen if we could create a synthetic neocortex that's substantially larger than the human version, say three times bigger.

Even more thought-provoking is his suggestion that we could connect our own biological neocortex to a much larger, synthetic one in the 'cloud'[114] (like the Internet today).

It almost goes without saying that Kurzweil believes AGI is not just achievable, but will be with us soon (he predicts the Turing Test will be passed in the late 2020s). And he paints a rosy picture of what life will look like when enriched by AGI, and all that comes with it.

Michio Kaku

Michio Kaku is a prolific writer, speaker and broadcaster on technology and the future. He has also done much to popularize science to the general public. He is an accomplished physicist and scientist, and his greatest academic accomplishment was co-founding 'string field theory', an advanced subset of theoretical physics.

Arguably, his greater gift is an ability to communicate complex ideas about technology.[115] His presentation style has been described as charismatic, and he is known for painting cogent pictures of the future technology could bring. His books are mostly about space, time and the human mind, and the best-known titles include *Physics of the Impossible*, *Physics of the Future* and *The Future of the Mind*.

While his books have sold well, he has reached a wider audience through radio, TV and film work. These include pieces for the BBC, Discovery Channel and Science Channel, as well as appearances on mainstream programmes for CNN, CBS, NBC, Fox News and Al

Jazeera. One of his books, *Hyperspace*, even inspired a platinum-selling rock album, *Origin of Symmetry* by British band Muse.

Kaku's views on AI rarely provide 'big ideas' to transform our understanding of the subject. That's not meant pejoratively, actually the opposite. What he does instead is bring to life the ideas and principles of AI, and turn them into riveting stories, explanations and questions. For example, he's made thoughtful observations on the evolution of intelligence and offered arbitral views on the famous disagreement between Mark Zuckerberg and Elon Musk on whether AI is a threat.[116]

Kaku is another optimist about what AI will do to our lives in the future. He has spoken and written extensively on the convergence of AI and the human brain, and articulately presents the opportunities that will offer. For example, he mentions immortality frequently, often in terms of infinite knowledge as well as physical existence of the body.

Richard Feynman

Nobel Laureate Richard Feynman was a renowned physicist of the mid-20th century, and has near-legendary status for many students, researchers and teachers. He introduced the concepts of quantum mechanics to many, and his work was influential in the development of quantum computing and nanotechnology. His undergraduate lectures are required for reading for physics students to this day, and he was one of the first scientists to popularize science.

He's included here because of one particular lecture, in 1985, when he was asked for his views on AGI.[117]

He was answering questions 'off the cuff', and AI was not one of his areas of research. And yet he encapsulated many of the essential points of AI, in particular general intelligence, and used simple analogies and illustrations very effectively. He was asked:

> Do you think there will ever be a machine that will think like human beings and be more intelligent than human beings?

His answer, in a nutshell, was no, machines won't think like human beings, but probably yes, they will be more intelligent. And then went on to expound the headline answers in a wide-ranging discussion of what 'intelligence' and 'thinking' mean, the analogy with natural versus artificial motion, why computers will inevitably outperform humans in what we now call narrow AI, and several other points that now form part of AI education.

Strictly speaking, there's nothing to be learned from Feynman's contribution to AI that you won't find elsewhere. But it's so elegantly put, it's worth reading just to appreciate how he articulated it.

Nick Bostrom

Nick Bostrom is a Professor at Oxford University, where he co-founded the Future of Humanity Institute. He, like the Institute he leads, is multi-disciplined in his research, bringing together experience from philosophy, physics, neuroscience, logic, and of course AI.

He freely concedes that his work may initially seem 'somewhat scattered', but on closer look, it's all united by a common theme which he calls 'macro strategy … the study of how long-term outcomes for humanity may be connected to present-day actions'.[118]

His work is fascinating, and much of it is accessible to non-specialists. The area you might find of particular interest, and which stands out from the work of many others, is how he examines the role of technology. Much of what he writes begins with his assertion that:

> It is plausible that the long-term outcomes for our civilization depend sensitively on how we handle the introduction of certain transformative capabilities. Machine intelligence, in particular, is a big focus.

What comes across from his work are some thought-provoking insights, and illumination of the connections between many of the topics we've explored.

THE FUTURE ACCORDING TO PROFESSIONAL FUTUROLOGISTS

Alvin Toffler

Toffler was another well-known name in the field of writing about the future, but unlike those mentioned so far, his background was not in science, technology or even academia. His student years were split between an English degree and political activism. And also meeting and marrying his wife, Heidi, with whom he collaborated on much of his work. As well as a best-selling author, he was a journalist, educator and advisor to international business and political leaders.[119]

His approach to writing was the author's equivalent of method acting. For example, to gain experience of writing about mass production and business management, he and his wife spent 5 years working in factories, where she became a union leader. The research for his best-seller *Future Shock* took another 5 years and is considered a business writing classic.

The main thesis of the book was that when societies change too fast, they experience a trauma that leads to disruption of previous norms, such as decision-making. He labelled this concept 'Future Shock', and highlighted the rapid development of new technology as one of its causes.

Unlike the other writers mentioned here, he wrote little about AI explicitly. However, the relevance of his work on AI is clearest in his 1980 classic *The Third Wave*. In it, he continues the theme of Future Shock, but in the context of 'waves' of evolution of human society. The first two corresponded to the Agrarian and Industrial Revolutions. The third, in his view, was driven by technological revolution. He described this wave in terms such as information technology, electronics, space travel and global communication.

The value of Toffler's writing for people interested in AI is the idea of how human society reacts to rapid changes brought about by technology, in particular a change that is too rapid to absorb naturally. What 'too rapid' means is, of course, a question for debate. Toffler offers views on the subject, but his contribution was the question and its underlying idea, rather than any specific answers.

Amy Webb

Amy Webb is one of today's newer breed of professional futurists, coming to the field from a different direction to those we've seen so far.[120] Like Toffler, she has a journalism background, and no formal science or technology training. But during her stints as technology reporter at the Wall Street Journal and Newsweek, she developed a strong understanding of not just technology, but how it affects business and society.

She also developed a reputation for being able to communicate the subject in meaningful, useful terms that connected with her readers. This was demonstrated with the success of her first book, *Data: A Love Story*, followed by several others. She has written and spoken extensively on AI, and appears frequently on panels, boards and commissions tasked with considering the impact of AI on the world.

A specific reason for looking at her work in the context of this book, is her recent writing on best- and worst-case scenarios for how AI will affect us over the next half century. To do this, she examines the large US and Chinese companies who wield such influence over today's AI work. These include Google, Amazon, Baidu and Alibaba.

These are the only two 'professional futurologists' whose work we're going to describe here. There are plenty of others in the same business nowadays, and many have equally interesting and thought-provoking things to say. It's just that for the very specific aims of this book, they aren't quite as directly relevant as those above. A few whose writing you might want to explore include: Youngsook Park, the South Korean futurist who applies herself to the impact of modern technologies on social and political life; John Vary, John Lewis' in-house futurologist (yes, that's a real job!); Nicola Millard, another corporate futurologist, this time at BT; and Bryan Appleyard, the author and journalist, known for, among other things, his views on experts he believes are being overly optimistic about the future.

That brings to an end our look at what others have to say about the future of AI. They represent a wide cross-section of perspectives and backgrounds, and between them offer many invaluable insights into how we should form our own views.

SO FAR, SO GOOD?

No one can accurately predict the future of AI, and what the world will look like as it develops and propagates. Anyone that tries is at best making informed guesses. Even the greatest experts only offer opinions, not guarantees.

With the knowledge and understanding that this book has provided, you're in a position to speculate for yourself and make suggestions as valid as any you'll read elsewhere. We can't know if any of the experts will be correct, nor whether more fanciful pictures of an AI-powered future will prove to be wrong. But reading about them with an informed perspective gives you a firmer basis for your own bets on how the world will turn out when AI is more powerful, prevalent and prescient.

Of course, the only way to definitively know what AI will do in the future is to wait for it to arrive and experience it. Until then, we're not equipped to objectively, rationally and accurately figure out what the future holds. But while we're waiting, we can look to history for suggestions.

What we can glean from the past is that we've repeatedly surprised ourselves with the pace of progress and degree of impact of new technology developments. Over the last few centuries, technology has changed lives beyond recognition several times, on most occasions within a generation or so. Examples include the agrarian and industrial revolutions, the ability to travel around the world in hours, ubiquitous power, pocket-sized electronics and so on. Technology has caused generations of children to live in versions of the world their parents or grandparents wouldn't recognize. Much of this technology has had downsides, often leading to difficult questions around control and risk. In a few instances it's created the possibility of existential threats to the human race.

But we're still here, still developing as a species, still using technology to improve lives, and still expecting each generation to enjoy a better quality of life than those that came before.

So, while there's little we can definitively know about what AI will look like in the future, history offers a few repeating patterns which may well be relevant, as we take our next step towards it:

> We're not particularly good at predicting how technology develops and impacts us.
>
> So we're constantly grappling with the potential implications, especially the negative ones.
>
> But we manage to muddle along, usually finding some way of coping with the issues.
>
> In the meantime, the overall quality of our lives continues to improve.
>
> And so far, we're still here!

References

1. Nilsson N. J. *Artificial intelligence: A new synthesis*. San Francisco, CA: Morgan Kaufmann; 1998.
   Russell S. J., Norvig P. *Artificial intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall; 2014.
   Luger G. F. *Artificial intelligence: Structures and strategies for complex problem solving*. Boston; Harlow: Pearson Education; 2008.
   Poole D. L., Mackworth A. K., Goebel R. *Computational intelligence: A logical approach*. New York; Oxford: Oxford University Press; 1998.
2. Goertzel T. The path to more general artificial intelligence. *J Exp Theor Artif In*. 2014; 26(3): 343–354.
3. Kim J. The problem of distinction between 'weak AI' and 'strong AI'. *J Society Philos Stud*. 2017; 117: 111.
4. Baron B., and State University Ball. In defense of strong AI. *Stance Int Undergrad Philos J*. 2017; 10: 38–49.
5. Bringsjord S. Strong AI is simply silly. *AI Magazine*. 15 March 1997; 18(1).
6. Searle J. R. Minds, brains, and programs. *Behav Brain Sci*. 2010; 3: 417–457.
7. Hodges A. Alan Turing: *The enigma*. Princeton, NJ: Princeton University Press; 2014.
8. Turing A. M. Computing machinery and intelligence. *Mind*. 1950; LIX(236): 433–460.
9. Wozniak S., Moon P. Three minutes with Steve Wozniak; 2007. Available from https://www.pcworld.com/article/134826/article.html (accessed 17 March 2020).
10. Goertzel B. What counts as a conscious thinking machine? *New Scientist Magazine*. 5 Sepetember 2012; (2881).
11. Nilsson N. J. Human-level artificial intelligence? Be serious! *AI Magazine*. 15 December 2005; 26(4).

12. Mayor A. *Gods and robots*. Princeton, NJ: Princeton University Press; 2018.
13. Strong J. S., Teiser S. *Relics of the Buddha*. Princeton, NJ: Princeton University Press; 2004.
14. Marshall V. Abu Mūsā Jābir Ibn Hayyān. *Royal Society of Chemistry: 175 Faces of Chemistry*; 2014. Available from http://www.rsc.org/diversity/175-faces/all-faces/abu-musa-jabir-ibn-hayyan/ (accessed 17 March 2020).
15. Thorstensen E. Creating golems: Uses of golem stories in the ethics of technologies. *NanoEthics*. 2017; 11(2): 153–168.
16. Shelley M. W. *Frankenstein*. New York, NY: Bantam Classics; 2003.
17. Butler S. *Darwin Among the Machines*. Christchurch, New Zealand: The Press; 1863.
18. Roberts A. *The palgrave history of science fiction*. London: Palgrave Macmillan; 2005.
19. Berggren J. L. *Episodes in the mathematics of medieval Islam*. Berlin: Springer Verlag; 1986.
20. Priani E. Ramon Llull. *The Stanford Encyclopedia of Philosophy*; Spring 2017. Available from https://plato.stanford.edu/archives/spr2017/entries/llull/ (accessed 17 March 2020).
21. Google. Google mission. Available from https://about.google/ (accessed 17 March 2020).
22. Davis M. *The universal computer: The road from Leibnitz to Turing*. New York, NY: W. W. Norton & Company; 2000.
23. Hobbes T. *Leviathan*. Scotts Valley: CreateSpace Independent Publishing Platform; 2011.
24. Leibniz G. L. *New essays concerning human understanding*. New York, NY: Macmillan; 1896.
25. Boole G. *An investigation of the laws of thought*. Gearhart, OR: Watchmaker Pub; 2010.
26. Frege G. *Begriffsschrift, Eine Der Arithmetischen Nachgebildete Formelsprache Des Reinen Denkens (Ed.1879)*. Hachette Livre-BNF; 2012.
27. Whitehead A. N., Russell B. *Principia mathematica*. Cambridge: Cambridge University Press; 1910.

28. Ewald W., Sieg W. *David Hilbert's lectures on the foundations of arithmetic and logic 1917–1933*. Berlin: Springer; 2013.
29. Gödel K. *On formally undecidable propositions of principia mathematica and related systems*. Courier Corporation; 1992.
30. Rahman W., Kurien P. *Blind men and the elephant: Demystifying the global IT services industry*. New Delhi; London: SAGE Publications; 2007.
31. The Project JMC Team. Professor John McCarthy, Father of AI. Available from: http://jmc.stanford.edu/ (accessed 17 March 2020).
32. McCarthy J., Minsky M. L., Rochester N., Shannon C. A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*. 2006; 27(4): 12–14.
33. Crevier D. *AI*. New York, NY: Basic Books; 1993.
34. Reavie V. Do you know the difference between data analytics and AI machine learning? *Forbes Agency Council* (2018). Available from https://www.forbes.com/sites/forbesagencycouncil/2018/08/01/do-you-know-the-difference-between-data-analytics-and-ai-machine-learning/#d8d64e15878d (accessed 17 March 2020).
35. 3RDi. AI turns enterprise search into intelligent search; 2019. Available from https://www.3rdisearch.com/ai-turns-enterprise-search-into-intelligent-search (accessed 17 March 2020).
36. Damron S. AI academy: What's the difference between forecasting and predictive modeling? 2017. Available from https://www.onemodel.co/blog/ai-academy-forecasting-vs-predictive-modeling (accessed 17 March 2020).
37. SciForce. Anomaly detection, a key task for AI and machine learning, explained; 2019. Available from https://www.kdnuggets.com/2019/10/anomaly-detection-explained.html (accessed 17 March 2020).
38. Venables M. An overview of computer vision. *Towards Data Science*; 2019. Available from https://towardsdatascience.com/an-overview-of-computer-vision-1f75c2ab1b66 (accessed 17 March 2020).
39. Tavares J. Image processing and analysis: Applications and trends. *AES-ATEMA International Conference Series - Advances and Trends in Engineering Materials and their Applications*; 2010.

40. Kacalak W, Stuart K. D., Majewski M. Selected problems of intelligent handwriting recognition in *Analysis and Design of Intelligent Systems Using Soft Computing Techniques*, ed. Patricia Melin et al.; Berlin, Heidelberg: Springer; 2007.
41. Shrivastava V. Video processing: The new paradigm in artificial intelligence. 2019. Available from https://medium.com/@AuthorVivek/video-processing-the-new-paradigm-in-artificial-intelligence-3cf7cb26a582 (accessed 17 March 2020).
42. Adobe. About Color; 2018. Available from https://helpx.adobe.com/uk/photoshop/using/color.html (accessed 17 March 2020).
43. Collobert R., Weston J., Bottou L., Karlen M., Kavukcuoglu K., Kuksa P. Natural language processing (almost) from scratch. *Computing Research Repository - CORR* 12; 2011.
44. Davenport M., Hannahs S. J. Phonemic analysis, in *Introducing Phonetics and Phonology* (Fourth edition). Abingdon, Oxon; New York, NY: Routledge, 2020. Previous edition: London: Hodder Education, {2010}: Routledge; 2020.
45. Kamath U., Liu J., Whitaker J. *Deep learning for NLP and speech recognition*. Cham: Springer International Publishing; 2019.
46. Bhatt R., Gupta P. Sentiment analysis. *Indian J Sci Technol*. 2019; 12(41): 1–6.
47. Ganesan K. All you need to know about text preprocessing for NLP and machine learning; 2019. Available from https://www.kdnuggets.com/2019/04/text-preprocessing-nlp-machine-learning.html (accessed 17 March 2020).
48. Piatetsky G. Data science, machine learning: Main developments in 2017 and key trends in 2018; 2017. Available from: https://www.kdnuggets.com/2017/12/data-science-machine-learning-main-developments-trends.html (accessed 17 March 2020).
49. Rouse M., Rosencrance L. What is RPA? Everything you need to know; 2020. Available from https://searchcio.techtarget.com/definition/RPA (accessed 17 March 2020).
50. Schmelzer R. Making the Internet of Things (IoT) more intelligent with AI; 2019. Available from https://www.forbes.com/sites/

cognitiveworld/2019/10/01/making-the-internet-of-things-iot-more-intelligent-with-ai/#57b47a28fd9b (accessed 17 March 2020).
51. Marr B. What is the difference between AI and robotics? 2019. Available from https://bernardmarr.com/default.asp?contentID <hig>=</hig>1908 (accessed 17 March 2020).
52. TWI. What is an autonomous vehicle? 2019. Available from https://www.twi-global.com/technical-knowledge/faqs/what-is-an-autonomous-vehicle (accessed 17 March 2020).
53. Talabis M., McPherson R., Miyamoto I., Martin J. Analytics defined in *Information Security Analytics*. Amsterdam: Elsevier; 2015.
54. Maini V. Machine learning for humans, Part 3: Unsupervised learning. *Towards Data Science*; 2017. Available from https://medium.com/machine-learning-for-humans/unsupervised-learning-f45587588294 (accessed 17 March 2020).
55. Osiński B., Budek K. What is reinforcement learning? The complete guide; 2018. Available from https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/ (accessed 17 March 2020).
56. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015; 521(7553): 436–434.
57. Schmidhuber J. Deep learning in neural networks: An overview. *Neur Net*. 2015; 6185–6117.
58. Svozil D., Kvasnicka V., Pospichal J. Introduction to multi-layer feed-forward neural networks. *Chemometr Int Lab*. 1997; 39(1): 43–62.
59. Sherstinsky A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D*. 2020; 404.
60. Ian G., Yoshua B., Aaron C. *Deep learning*. Cambridge, MA: MIT Press; 2016.
61. Khan A. R., Asgar M. Z. An intelligent agent for a vacuum cleaner. *Int J Digit Cont*. 2009; 3(2): 143–146.
62. Webb D. Humans V. Robots: The best movies and shows about AI; 2020. Available from https://filmdaily.co/obsessions/best-movies-and-shows-about-ai/ (accessed 17 March 2020).

63. Dennys H. Why sushi boss may say yo! To robot waiters: It's just one idea on a bold menu to revive the restaurant sector; 2019. Available from https://www.thisismoney.co.uk/money/markets/article-7489625/Why-sushi-boss-say-Yo-robot-waiters.html (accessed 17 March 2020).
64. Feldberg S. Robot Butlers deliver snacks and delight guests at the Vdara; 2018. Available from https://www.travelweekly.com/North-America-Travel/Robot-butlers-deliver-snacks-and-delight-guests-at-the-Vdara (accessed 17 March 2020).
65. Honda. Honda Robotics; 2020. Available from https://global.honda/innovation/robotics/ASIMO.html (accessed 17 March 2020).
66. Hyundai Robotics. About Hyundai Robotics; 2020. Available from https://www.hyundai-robotics.com/english/intro/intro3.html (accessed 17 March 2020).
67. Asus. Zenbo; 2020. Available from https://www.asus.com/Commercial-Intelligent-Robot/Zenbo/ (accessed 17 March 2020).
68. Samsung Research. Robot; 2020. Available from https://research.samsung.com/robot (accessed 17 March 2020).
69. Hanson Robotics. Sophia; 2018. Available from https://www.hansonrobotics.com/sophia/ (accessed 17 March 2020).
70. Wang X., Wang Y., Xu C., Ke Y. A neural network GPS navigation algorithm. *Journal of Beijing Institute of Technology 5*. 1998; 18: 1–3.
71. Krzanich B. Data is the new oil in the future of automated driving; 2016. Available from https://newsroom.intel.com/editorials/krzanich-the-future-of-automated-driving (accessed 29 June 2020).
72. Du K., Zhu F., Chou C. C. A new data-driven design method for thin-walled vehicular structures under crash loading. *SAE Int J Trans Saf*. 2017; 5(2): 188–193.
73. Huang S.-S., Hsiao P.-Y. Occupant classification for smart airbag using Bayesian Filtering. *2010 International Conference on Green Circuits and Systems*. 2010; IEEE: 660–665.
74. Somasundaram M., Angela E. M., Thilagavathy K., Soundharya R. AI enabled blind spot detection using RCNN based image processing. *International Journal of Recent Technology and Engineering (IJRTE)*. 2019; 8(2S5): 28–30.

75. Gao R. Uber now has a panic button in the US; 2018. Available from https://www.androidpolice.com/2018/05/29/uber-now-panic-button-us/ (accessed 17 March 2020).
76. Kashyaap S., Putrevu S. What is playing inside Ola Play; 2018. Available from https://yourstory.com/2018/05/what-is-playing-inside-ola-play (accessed 17 March 2020).
77. Ackert S. Engine maintenance concepts for financiers. *Aircraft Monitor*. 2011; 2: 1–43.
78. Kashyap R. Artificial intelligence systems in aviation in *Advances in Computer and Electrical Engineering: Cases on Modern Computer Systems in Aviation*. Pennsylvania: IGI Global; 2019.
79. Karna K. N. Artificial intelligence in intelligent transportation systems. *I V H S J*. 1995; 2(3): iii–viii.
80. Scheffer C. Train scheduling: Hardness and algorithms in *Walcom: Algorithms and Computation: Lecture Notes in Computer Science*. Cham: Springer International Publishing; 2020.
81. Porter, M. E. The value chain and competitive advantage in *Understanding Business Processes*, ed. David Barnes, 50–66. London: Routledge; 2001.
82. Hall J. How artificial intelligence is transforming digital marketing; 2019. Available from https://www.forbes.com/sites/forbesagencycouncil/2019/08/21/how-artificial-intelligence-is-transforming-digital-marketing/#2dd5a60121e1 (accessed 17 March 2020).
83. Digitaland. Why advertisers are turning to AI to build smarter banner ads; 2017. Available from https://www.digitaland.tv/blog/advertisers-turning-ai-build-smarter-banner-ads/ (accessed 17 March 2020).
84. Miles S. How retailers use AI, mapping to boost ROI on store remodels; 2019. Available from https://streetfightmag.com/2019/09/13/how-retailers-use-ai-mapping-to-boost-roi-on-store-remodels#.XnIz_pP7QWo (accessed 17 March 2020).
85. Fashion Network. Fashion portal Rent it Bae launches first AI retail store; 2018. Available from https://in.fashionnetwork.com/news/Fashion-portal-rent-it-bae-launches-first-ai-retail-store,1012323.html (accessed 17 March 2020).

86. Lowe's Innovation Labs. Robotics; 2018. Available from http://www.lowesinnovationlabs.com/robotics (accessed 17 March 2020).
87. Shepherd A. The evolution continues: The lush labs app; 2019. Availabe from https://uk.lush.com/article/evolution-continues-lush-labs-app (accessed 17 March 2020).
88. L. G. Sephora and artificial intelligence: What does the future of beauty look like? 2018. Available from https://digital.hbs.edu/platform-rctom/submission/sephora-and-artificial-intelligence-what-does-the-future-of-beauty-look-like/ (accessed 17 March 2020).
89. Lenskart. 3d Try on Eyewear; 2020. Available from https://www.lenskart.com/compare-looks (accessed 17 March 2020).
90. Grill-Goodman J. First look: The technology behind Neiman Marcus's new retail experience at New York's Hudson Yards; 2019. Available from https://risnews.com/first-look-technology-behind-neiman-marcuss-new-retail-experience-new-yorks-hudson-yards (accessed 17 March 2020).
91. Sudha C., Nirmal Raj T. Analysis of suspicious pattern discovery using AI-neural network in credit card fraud detection. *Int J Cur Res Rev*. 2017; 9(10).
92. Obukhov A., Krasnyanskiy M., Nikolyukin M. Implementation of decision support subsystem in electronic document systems using machine learning techniques. 2019 International Multi-Conference on Industrial Engineering and Modern Technologies. Fareastcon; 2019.
93. Rouse, Rosencrance. What is RPA? Everything you need to know; 2015. Available from https://searchcio.techtarget.com/definition/RPA (accessed 29 June 2020).
94. Berthene A. REI uses artificial intelligence for order management; 2019. Available from https://www.digitalcommerce360.com/2019/03/11/rei-uses-artificial-intelligence-for-order-management/ (accessed 17 March 2020).
95. OMQ. Omq reply: Automatically reply to e-Mails with AI; 2020. Available from https://www.omq.ai/products/reply/ (accessed 17 March 2020).
96. KeyPay. Is artificial intelligence the future for payroll?; 2019. Available from https://www.accountingweb.co.uk/community/

industry-insights/is-artificial-intelligence-the-future-for-payroll (accessed 17 March 2020).
97. Centre for Data Ethics and Innovation. AI and personal insurance. *CDEI Snapshot Series*; 2019. Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/833203/AI_and_Insurance_-_WEB.pdf (accessed 17 March 2020).
98. Kuprenko V. Artificial intelligence in the logistics industry - the network effect; 2019. Available from https://supplychainbeyond.com/artificial-intelligence-in-the-logistics-industry/ (accessed 17 March 2020).
99. Read S. Argos launches same day delivery in first for industry. *Evening Standard*; 7 October 2015. Available from https://www.standard.co.uk/business/argos-launches-same-day-delivery-in-first-for-industry-a3084411 (accessed 29 June 2020).
100. Rahman W. Starbucks isn't a coffee business—it's a data tech company. *Medium*; 2020. Available from https://marker.medium.com/starbucks-isnt-a-coffee-company-its-a-data-technology-business-ddd9b397d83e (accessed 17 March 2020).
101. Makortoff K. Barclays using 'Big Brother' tactics to spy on staff, says Tuc. *The Guardian*; 20 February 2020. Available from https://www.theguardian.com/business/2020/feb/20/barlays-using-dytopian-big-brother-tactics-to-spy-on-staff-says-tuc (accessed 29 June 2020).
102. Thomson J. J. Killing, letting die, and the trolley problem. *Monist*. 1976; 59(2): 214–217.
103. Wareham M. Campaign to stop killer robots; 2020. Available from https://www.stopkillerrobots.org/ (accessed 17 March 2020).
104. McKendrick J. Nine companies are shaping the future of artificial intelligence; 2019. Available from https://www.forbes.com/sites/joemckendrick/2019/04/10/nine-companies-are-shaping-the-future-of-artificial-intelligence/#2e59650f2cf1 (accessed 17 March 2020).
105. Holford N. The big three - Asimov, Clarke and Heinlein; 2019. Available from https://www.sfandfantasy.co.uk/php/the-big-3.php (accessed 17 March 2020).

106. Gabrynowicz J. I. Haley archive: Arthur C. Clarke August 1956 Letter on GPS, DBS, and More. *Res Communis (University of Mississippi)*; 2008. Available from https://rescommunis.wordpress.com/2008/03/27/haley-archive-arthur-c-clarke-august-1956-letter-on-gps-dbs-and-more/ (accessed 17 March 2020).
107. UNESCO. Kalinga Prize Laureate 1961 Arthur C. Clarke; 2020. Available from http://www.unesco.org/new/en/natural-sciences/science-technology/sti-policy/global-focus/science-popularization/prizes/kalinga-prize/kalinga-winners/kalinga-1961/ (accessed 17 March 2020).
108. Chiasson D. 2001: A Space Odyssey: What it Means, and How it Was Made. *The New Yorker*. 2018.
109. Marshall C. How Isaac Asimov went from Star Trek Critic to Star Trek Fan & Advisor; 2015. Available from http://www.openculture.com/2015/12/how-isaac-asimov-went-from-star-trek-critic-to-star-trek-fan-advisor.html (accessed 17 March 2020).
110. Sheppard S. Asimov's 'Robotics' FAQ; 2000. Available from http://www.heretical.com/miscella/asimov.html (accessed 17 March 2020).
111. Sheppard. Asimov's 'Robotics' FAQ.
112. P. B. S. Ray Kurzweil; 2004. Available from https://www.pbs.org/wgbh/theymadeamerica/whomade/kurzweil_hi.html (accessed 17 March 2020).
113. Reedy C. Kurzweil claims that the singularity will happen by 2045; 2017. Available from https://futurism.com/kurzweil-claims-that-the-singularity-will-happen-by-2045 (accessed 17 March 2020).
114. Israel S. Artificial intelligence, human brain to merge in 2030s, says Futurist Kurzweil; 2015. Available from https://www.kurzweilai.net/cbc-artificial-intelligence-human-brain-to-merge-in-2030s-says-futurist-kurzweil (accessed 17 March 2020).
115. Kaku M. Michio Kaku picks five books to help you understand the future. *The Guardian*; 2018. Available from https://www.theguardian.com/books/2018/mar/12/further-reading-michio-kaku-books-to-understand-future (accessed 17 March 2020).

116. Lee L. Who is correct? Elon Musk or Mark Zuckerberg?; 2017. Available from http://www.enterpriseitnews.com.my/who-is-correct-elon-musk-or-mark-zuckerberg/ (accessed 17 March 2020).
117. Veisdal V. Richard Feynman on artificial general intelligence. *Cantor's Paradise*; 2019. Available from https://cantorsparadise.substack.com/p/richard-feynman-on-artificial-general (accessed 17 March 2020).
118. Bostrom N. Bio. *Nick Bostrom's Home Page*; 2020. Available from https://nickbostrom.com/#bio (accessed 17 March 2020).
119. Toffler Associates. The Toffler Legacy; 2020. Available from https://www.tofflerassociates.com/about/the-toffler-legacy/ (accessed 17 March 2020).
120. Cooney C. Amy Webb, Founder. *Future Today Institute*; 2020. Available from https://futuretodayinstitute.com/amy-webb/ (accessed 17 March 2020).

About the Author

Was Rahman is the CEO of AI Prescience, an AI research and consulting firm. He has over 30 years of global experience using data and technology to improve business performance, particularly AI.

Was writes extensively on technology in business, for both mainstream and academic audiences. His work can be found at Amazon, Medium and ORCID.

His research interests include the impact of AI on organizational decision-making, effective AI ethics and governance, and the role of AI in new forms of social division.

In business, he has worked with large corporates, start-ups and SMEs around the world, advising CEOs, Boards and Investors. He has held leadership roles at Accenture, Infosys and Wipro, managing business in the United States, European Union and Asia Pacific. He has also run start-ups and raised funding.

He has also worked for various governments: Was has briefed UK and Indian political leaders and officials on technology industry policy. In 2008, he was appointed sector specialist for the UK Government, advising ministers and their teams on the global technology industry and establishing UK–India technology partnerships.

He graduated in physics from Oxford, then studied computer science for master's degree at Coventry. His AI and data science education is courtesy of Stanford, Johns Hopkins, Amazon and Google, among others. He has been a guest lecturer at Oxford's

Saïd Business School, the Judge Business School at Cambridge, London Business School and IIT Madras.

An experienced public speaker, he has chaired and participated in numerous international conference sessions. He has been invited to speak at events organized by various institutions, including the IoD, NASSCOM, the *Financial Times* and the Governments of India and the United Kingdom. He has participated in, and supported, many high-profile international events, including the inaugural World Economic Forum 'Young Global Leaders' meeting, and the UK Prime Minister's launch of London's Tech City.

Was splits his time between London (United Kingdom) and Bangalore (India). Outside work, he is obsessive about coffee, photography, classic cars and music.