palgrave pivot

Artificial Intelligence for Business

Innovation, Tools and Practices

Edited by

Ana Landeta Echeberria

palgrave macmillan

Artificial Intelligence for Business

Ana Landeta Echeberria
Editor

Artificial Intelligence for Business

Innovation, Tools and Practices

palgrave
macmillan

Editor
Ana Landeta Echeberria
Faculty of Economics and Business
Universidad a Distancia de Madrid (UDIMA)
Madrid, Spain

ISBN 978-3-030-88240-2 ISBN 978-3-030-88241-9 (eBook)
https://doi.org/10.1007/978-3-030-88241-9

© The Editor(s) (if applicable) and The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover pattern © Harvey Loake

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Dear readers,

The book builds on the previous research and real business practices of Universidad a Distancia de Madrid (Spain), "Technological Social Sciences Research Group" and TodoStartups Ltd. to draw up principles to facilitate innovation, adoption and trust in Artificial Intelligence (AI). Their debates inspired some of the principles of how AI is currently used and its future potential in day-to-day areas of application.

Moreover, this work emphasizes the need to adopt real AI practices and also maps the economic, legal and social impacts of the applications and implications of AI technologies, presenting real AI practices.

AI as a field is still in its formative stages, and the diversity of approaches and applicability methods and techniques are crucial to its development. Until now, AI theory is mainly based on techniques and methods for the technical purposes of ICT. Therefore, this book constitutes an academic-professional guide, "A how to use AI techniques and methods for real purposes", for AI performance in global scenarios.

This book takes current scholarship forward in its engagement with AI theory and practice for enterprises and applied research and innovation. Its main objectives are to:

- Evaluate the further implementation of AI business applications in all industrial sectors;
- Empirically explore the implications of AI evolution and impact in society, economics, labour market and finance;

- Examine the relationships between AI methods, techniques and technologies and
- Contribute to a better understanding of the challenge for the world-wide industry and to provide organizations with an opportunity to integrate AI within their daily activities.

AI is reshaping economies and promises to boost productivity and improve efficiency. These technologies, however, are still in an emergency phase, but they are ready to address global challenges and promote innovation and social benefits.

I hope you find it a useful and gratifying read.

Madrid, Spain
July 2021

Ana Landeta Echeberria

Overview

This book seeks to build a shared understanding of Artificial Intelligence (AI) today and in the near term within the global business scenario. It also seeks to co-ordinate and generate consistent discussions on this topic between different disciplines.

The book builds on the previous research and real business practices of Universidad a Distancia de Madrid (Spain), “Technological Social Sciences Research Group” and TodoStartups Ltd. to draw up principles to facilitate innovation, adoption and trust in Artificial Intelligence (AI). Their debates inspired some of the principles of how AI is currently used and its future potential in day-to-day areas of application.

This book takes current scholarship forward in its engagement with AI theory and practice for enterprises and applied research and innovation. Its main objectives are to:

- Evaluate the further implementation of AI business applications in all industrial sectors.
- Empirically explore the implications of AI evolution and impact in society, economics, labour market and finance,
- Examine the relationships between AI methods, techniques and technologies, and
- Contribute to a better understanding of the challenge for the world-wide industry and to provide organizations with an opportunity to integrate AI within their daily activities.

Therefore, it outlines international practices for the promotion of reliable AI systems, trends, research and development, fostering a digital ecosystem for AI and preparing companies for job transformation and building skills.

Acknowledgements

I would like to express my most sincere thanks to my colleagues and lecturers at Universidad a Distancia de Madrid (UDIMA): Dr María Aránzazu de las Heras García, Dr Juan Pazos Sierra, Dr David Lizcano Casas and Dr José Antonio Álvarez López for their dedication as co-authors during the drafting of this book and for the time they have so generously given, as well as their knowledge, know-how and friendship.

Without them, this book would not have seen the light of day.

I would also like to extend my thanks to the publishers who have believed in me and have greenlit this work. Special mention and thanks to Mrs Jessica Harrison, Mr Ruby Panigrahi and Mr Alec Selwyn for having patiently, and with the utmost dedication, guided me during the editing process.

And finally, my thanks to the business world, for being a great source of inspiration that lets us document and achieve results such as the ones presented in this book.

Contents

Notes on Contributors

José Antonio Álvarez López is a professor at the Universidad Rey Juan Carlos and Exes, a specialized IT training centre, both in Spain. He is the CEO in several companies and sectors, such as education and tax consulting. His research areas are related to disruptive innovation. He also participates in several business associations and business centres, such as Confederación de Empresarios de Madrid (CEIM), Club Financiero Genova, Droniberia, Financiero Genova, Droniberia, Spanish Distance Learning Association (ANCED) and so on.

David Lizcano Casas holds a PhD with Honors in Computer Science (2010) from the Universidad Politécnica de Madrid and an MSc Degree with Honors in Research in Complex Software Development (2008) from the Universidad Politécnica de Madrid, Spain. He is a professor and researcher in the Department of Computer Science at the Universidad a Distancia de Madrid (UDIMA). He held a research grant from the European Social Fund under their Research Personnel Training programme. He is involved in several national- and European-funded projects relating to Web 2.0 and Enterprise 2.0 technologies and service-oriented architectures (SOA), paradigms of programming, software engineering, human-computer interaction, computer-assisted E-learning and end-user development and programming.

María Aránzazu de las Heras García holds a PhD in Labor and Social Security Law and is a professor at the Universidad a Distancia de Madrid, Spain. Her research areas are the transformation of the world of work and the reconciliation of family and work life. She participates in professional associations and support for female leadership.

Ana Landeta Echeberria is a professor at Universidad a Distancia de Madrid, Spain, and an expert in corporate strategy and digital transformation. She holds a PhD in Management and Business Administration and serves as the Director of Institutional Relations and Director of "Technological Social Sciences" Research Group at the Universidad a Distancia de Madrid. She is also the Vice President of TodoStartups, S.L.

Juan Pazos Sierra received the first Spanish PhD Degree in Computer Science from the University Politécnica de Madrid (UPM), with a dissertation about artificial intelligence, where he was a Full Professor in the Department of Artificial Intelligence from 1975 until 2013. He is a Full Professor in Universidad a Distancia de Madrid. He also set up the First Spanish Artificial Intelligence Laboratory. He was a visiting professor at Carnegie Mellon University (CMU), among other universities. He is a member of the editorial board of the following journals: *AI Magazine*, *Expert Systems with Applications*, *Heuristics* and *Failure and Lessons Learned in Information Technology Management*, among others. He is the author or co-author of ten books on artificial intelligence and knowledge engineering, and 100 papers in ISI journals. His current Investigation & Development (I&D) is about technological singularity and artificial intelligence in Kardashev civilizations.

List of Figures

List of Tables

Introduction

This book seeks to build a shared understanding of Artificial Intelligence (AI) today and in the near term within the global business scenario. It maps the economic, legal and social impacts of the application of AI technologies and their implications, presenting evidence and real-life case studies. It also seeks to co-ordinate and generate consistent discussions on this topic between different disciplines.

Therefore, it outlines international practices for the promotion of reliable AI systems, trends, research and development, fostering a digital ecosystem for AI and preparing companies for job transformation and building skills.

Chapter 1, "The Foundations of AI", provides a historical overview of AI evolution from its symbolic development—AI definition, history, fundamental framework and main elements. It also proposes a general framework for comprehending AI for business purposes. Chapter 2, "AI: Methods and Techniques. Knowledge-Based Systems", reviews expert systems (production systems, structure and components of rule-based systems); modes of reasoning and knowledge acquisition; knowledge-based systems (artificial neural networks and their definition, typology, construction methodology and learning). Chapter 3, "The Impact of AI in Business, Economics and Innovation", offers guidance from consultants and identifies documented professional practices with reference to the current overall impact of Artificial Intelligence (AI) in business, economics and innovation. Consequently, it focuses on business-oriented design, AI tools to model business processes and the benefits of AI technologies. It also shows how business leaders can remain competitive in the new economic environment, developing the required skills to understand the economic implications of AI, considering the changes that business must undertake in order to address the economic and social implications of large-scale AI applications. Additionally, it highlights the importance, benefits and applications of Machine Learning and proposes a future AI research agenda for certain industries (Strategy, Relationship Marketing, Servicescape, Customer acceptance, Social acceptance, Management, Workforce and Transhumanism).

Chapter 4, "AI Implications for the Future of Work", illustrates key concepts in understanding the implications of AI for employment in the future; digital transformation in the workplace; AI for work-related decision-making; the impact of robotisation processes on the labour market; new ways of working (teleworking and working on digital platforms); new professions in the Industry 4.0 and the impact of AI on education.

Chapter 5, "AI Integration in the Digital Transformation Strategy", presents three frameworks (Digital Strategic Framework, AI readiness and AI integration in Digital Transformation Strategy Model) that help to address new businesses challenges and a series of strategic-operational suggestions that permit an initial and future approach associated with AI integration, as new technological elements of the digital transformation strategy.

Chapter 6 presents several "Cases of AI Use and Applications in Every Industry and Business Functions", assessing and implementing business applications: Customer service, Consumer: Marketing and sales, Energy, resources and industries, Financial services and FinTech, Government and Public Administration, Life Sciences and healthcare, HealthTech, Retail, Mobility and self-driving cars, Human Resources, Operations and Startups.

Finally, Chap. 7, "Technological Singularity and Ethical Issues of AI", explores the possibility of achieving an artificial general intelligence that exceeds human intelligence, one of the major paradigms of life for humanity today according to certain current ethical issues of AI.

Ana Landeta Echeberria

CHAPTER 1

The Foundations of AI

David Lizcano Casas and Juan Pazos Sierra

Abstract From its most remote past, humanity has always sought to build artefacts that emulate or at least simulate certain aspects of intelligent beings and especially intelligence. As a matter of fact all myths and realisations concerning automatons and homunculi are but the expression of an ineluctable desire of human beings to surpass themselves in all walks of life and especially in the area of intelligence. This chapter delves into the definition of artificial intelligence, its history and evolution, the elements that can shape AI theory and the current frame of reference in the field.

Keywords AI • AI theory • AI history • Framework for AI

1.1 Defining AI

When Solomon said: "there is no new thing under the sun", he was attesting to something that has been repeated numerous times throughout history in all walks of life, and especially in science. Indeed, any attempt to

D. L. Casas (✉) • J. P. Sierra
Universidad a Distancia de Madrid (UDIMA), Madrid, Spain
e-mail: david.lizcano@udima.es; juan.pazos@udima.es

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_1

research the precise meaning of the elusive word intelligence drags us into a semantic labyrinth that arises from the various uses of the term. Thus, if we seek recourse to dictionaries in an attempt to define intelligence and, for example, look at *Webster's New World Dictionary* (Guralnik, 1969), we find the following two definitions: (a) The ability to learn or understand from experience; ability to acquire and retain knowledge; mental ability and (b) the ability to respond quickly and successfully to a new situation; use of the faculty of reason in solving problems and so on effectively. In this regard, Minsky defined intelligence as follows: It is only a word that people use to name those unknown processes with which our brains solve problems we call hard. In other words, it is to solve problems that have not yet been understood, because once we know how to solve them, they are no longer seen as something that requires intelligence. Tesler expressed something similar, albeit ironically, when he said: Once a mental function is programmed, people no longer consider it an essential ingredient of real thought. In this regard, the unavoidable essence of intelligence is always in that next thing that hasn't been programmed yet. In other words, artificial intelligence (AI) is all that hasn't been done yet.

Russel and Norvig look at the following definitions (Russell and Norvig, 2005):

1. The interesting task of making computers think … machines with minds, in its wide literary sense, by Haugeland.
2. [The automation of] activities that we link to human thought processes, activities such as decision-making, problem-solving, learning…, by Bellman.
3. The art of creating machines that are able to perform functions that, when performed by people, require intelligence, by Kurzweil.
4. The study of how to make computers perform tasks that, for now, human perform better, by Rich and Knight.
5. The study of mental faculties by using computational models, by Charniak and McDermott.
6. The study of the calculations that enable perception, reasoning and action, by Winston.
7. A field of study that centres on the explanation and emulation of intelligent behaviour according to computational processes, by Schalkoff.
8. And, the branch of computer science that deals with the automation of intelligent conduct, by Luger and Stubblefield.

These definitions vary with regard to two main dimensions. Definitions 1, 2, 5 and 6 refer to mental processes and reasoning, whereas 3, 4, 7 and 8 refer to conduct. However, 1, 2, 3 and 4 measure the desirable condition according to human efficiency, while 5, 6, 7 and 8 do so according to a concept of ideal intelligence, called rationality, given that a system is considered rational if it does the correct thing. Thus, according to the above, there are four possible goals to be achieved in AI, namely: Systems that think like human beings, systems that think rationally and systems that act like human beings or systems that act rationally.

As is evident, these definitions are barely workable, due to their imprecision and fuzziness. Nevertheless, as Lwoff, the winner of the Nobel Prize for Medicine, points out, defining is one of the means to discovery. As a matter of fact, it is an excellent heuristic method, as it forces us to boil down the essentials of a category or a phenomenon in a formula, so that contains all that it needs to contain and excludes all that must be excluded. Therefore, it is worth framing a good definition, as this exercise forces us to make a critical consideration of all the terms or aspects of a problem. What Lwoff does not say, is that there are two difficulties to establishing a good definition: The first is intrinsic to the search for an ideal definition, and is directly linked to the philosophical problem of nominalism. The second, of a more practical nature, is no less tricky. It is that dictionaries are self-contained, that is to say, all the words that exist in them are expressed in terms of those same words, which mean, sooner or later, a word will appear that has already appeared before and as the defined cannot be a part of the definition, we are faced with a contradiction. Still, and in spite of all the difficulties that it entails, it is necessary to define intelligence, at least in the sense that Kant means it: A definition does not only replace the name of a thing for other, more comprehensible words, but it also contains a clear distinction that enables the certain knowledge of the defined object and helps to apply the explained concept.

Keeping all of the above in mind, we are going to provide an injunction, in Hessenstein's words, by listing certain constituting properties of intelligence (Borrajo et al., 1993). These include the following abilities: Learning, abductive, deductive and inductive reasoning. Analysis, or to discriminate and categorise or identify the relative importance of the different elements of a situation. Synthesis, to find similarities between several situations, in spite of any differences that set them apart. To discover differences between several situations, despite any similarities that connect them. Flexibility and adaptability to respond to concrete situations. To synthesise new concepts from basic old concepts taken and rearranged in other ways. To discover order in apparently chaotic situations. To make sense of ambiguous and contradictory messages. To come up with new ideas. To solve problems and make the most convenient decision in each situation. To foresee and prospect the future. To continuously gather knowledge from all available sources and assemble it as an integrated and congruent whole. To use all accessible and available information and add apparently unrelated items of information to create new and significant perspectives. To always be innovative and creative. To correctly establish priorities. To understand situations beyond what is superficially apparent, that is, beneath the symptoms. To identify the objects behind situations. To restate and solve problems. To consider the consequences of actions. To anticipate future developments and resulting plans. Put in the effort proportional to the situation. To exhibit coherent and desirable behaviour depending on each specific situation. To provoke and to detect serendipity.

Once the injunction of the elusive noun “intelligence” has been established, the ambiguous adjective “artificial” shall be considered, taking into account the following words pronounced by Robert Sokolowski in 1988:

> This adjective can be used in two senses, and it is important to determine which one applies in the term artificial intelligence. The word artificial is used in one sense when it is applied, say, to flowers and in other sense when it is applied to light. In both cases something is called artificial because it is fabricated. But in the first usage artificial means that the thing seems to be, but really is not, what it looks like. The artificial is the merely apparent; it just shows how something else looks. Artificial flowers are only paper, not flowers at all; anyone who takes them to be flowers is mistaken. But artificial light is light and it does illuminate. It is fabricated as a substitute for natural light, but once fabricated it is what it seems to be. In this sense the artificial is not the merely apparent, not simply an imitation of something else. The appearance of the thing reveals what it is, not how something else looks.

Obviously, here and now, we shall use the second meaning of the adjective.

1.2 Artificial Intelligence: History

Seen retrospectively, AI has passed through the following stages since its birth:

1.2.1 E1: The Secret Period (1932–1955)

From its most remote past, humanity has always sought to build artefacts that emulate or at least simulate certain aspects of intelligent beings and especially intelligence. As a matter of fact all myths and realisations concerning automatons and homunculi are but the expression of an ineluctable desire of human beings to surpass themselves in all walks of life and especially, in the area of intelligence. We witness this in classical Greek literature, more specifically, in Homer.

In the Iliad, when Thetis visits the god Hephaestus's workshop to commission a set of armour for her son Achilles, it is mentioned that the disabled god has built two female androids. These robots were made of solid gold, possessed intelligence in their minds, had the ability to speak and in their role as loyal attendants helped the god to walk on his deformed legs, and when their work was done, they would return to their inanimate state. Although many attempts were made to build androids of all types, none of them possessed any intelligence.

Years later, in 1887, Peirce wrote the following about thought outside of its natural habitat: The human brain: Precisely how much the business of thinking a machine could possibly be made to perform, and what part of it must be left to the living mind is a question not without conceivable practical importance.

The imaginative inventor and Lucasian professor of mathematics, Babbage, stated that his analytical machine would be able to play games like chess, in those times, the epitome of a problem requiring intelligence. As a matter of fact, Babbage's desire was to build machines that could think, learn and believe and whose ability to perform these functions would grow until the category of the problems solved by them grew to encompass the problems to which human beings apply their minds. Unfortunately, Babbage was unable to build his machine.

It was not until 1932 when the first real work that included a heuristic technique widely used in AI today was performed. It was performed by Rejewski, Eozyeahi and Zygalski at the Biuro Szyfrów, the Polish cypher bureau, in order to decode the messages coded by the German Enigma machine (Hodges, 1985; Teuscher, 2004). After analysing German messages encrypted with the same key, comparing the tables generated from messages on same and different days, that is, with similar and different encryption keys respectively, Rejewski and his colleagues came to the following two basic conclusions: One, the tables varied when the key did, which meant there was a connection between them. And two, the number of connections was not modified regardless of the configuration of the plugboard which exchanged pairs of letters and which constituted the main security element of the Enigma machine, making them virtually unbreakable. This fundamental discovery meant that Rejewski and his colleagues no longer had to find the correct encryption key from several thousand billion combinations, but only needed to guess the initial position of the rotors from 100,000 possible positions. This combinatorial explosion was reduced thanks to the heuristic “ignore the commutator”, it being an element placed between the keyboard and the first modifier in Enigma. This heuristic is widely used in the AI heuristic procedure of generating and testing. In order to perform this tedious task, the Polish cryptanalysts built a machine they called the *bomba*. According to one version, it was because Rejewski hit upon the solution of how to decipher Enigma while eating an ice cream in the shape of a half sphere called *bomba*. This continued until 31 December 1938 when the Nazi cryptographers increased Enigma’s security to the point that the number of possible keys increased to 159,000,000,000,000,000,000. Thereafter, due to the lack of material resources, Rejewski’s group felt unable to continue deciphering the messages of the new Enigma machine. Their chief, Commander Lager, contacted his French and British counterparts on 30 June 1939 for a work meeting in Warsaw on 24 July. This meeting left the British and the French astounded with the progress made by the Polish in cryptography. As a present, Langer gave his colleagues several replicas of the Enigma machine and plans to build *bombas*, which were sent to Paris in diplomatic pouches. From there, on 16 August, one of these machines made its way to London. To avoid arousing the suspicion of Nazi spies, it was concealed in the baggage of the playwright Sacha Guitry and his wife, the actress Yvonne Printemps.

The Polish advancements had demonstrated to their allies the importance of using mathematicians, scientists, expert chess players and crossword addicts; thus they were recruited to join forces with linguists and classical scholars who until then, had been tasked with the decryption. It was the first time that a multidisciplinary team had been used to solve a problem, ahead of the Operations Research teams.

They were all taken to the headquarters of the Government Code and Cipher School, GC&CS, newly established in Bletchley Park in Buckinghamshire, some 80 km from London, headed by Commander Denniston.

Among the recruited mathematicians was Turing (Anguera de Sojo et al., 2020; Lara et al., 2021) who had previously published a paper in 1936 on two types of machines, namely: One, the automatic machine or a-machine, equivalent to an algorithm and which, in his honour, his PhD supervisor called the Turing Machine, and its generalisation, called the Universal a-machine, UTM. And two, the c-machine, or choice machine. While the first type dealt with the question of algorithmic computability, the second type demonstrated that computation went beyond algorithms, given that the c-machine interacted with an operator, that is, a human user. This means that from the very beginning, Turing had been aware that algorithmic computation was not the only possibility.

Two years later, in 1938, in his PhD thesis, he presented a new type of machine, the oracle machine or o-machine. Formally, an oracle may be described as a set that may be queried on any value and returns "true" if the requested value is in the set and "false" if it is not. Specifically, Turing excluded the possibility that an oracle could be an effective computational entity, saying: We shall not go any further into the nature of the oracle apart from saying that it cannot be a machine. That is to say, Turing's oracles were used to represent non-computable information obtained outside the system.

Three years later, in 1941, Turing was already thinking about intelligent machines and more concretely, about the possibility of building computing machines that could solve problems by searching through a space of possible solutions guided by heuristic rules or rules of thumb, and on the mechanisation of chess. In his leisure time at Bletchley Park, he discussed these topics and machine learning with his colleagues. As a matter of fact, he distributed a typewritten essay concerning intelligent machines, as attested to by Michie, another of his colleagues in Bletchley, which was unfortunately lost. This was undoubtedly the first article on AI.

Nevertheless, the important task for which he was hired was to continue the work of the Polish cryptanalysts and decipher the new Enigma machines, which were more secure and difficult to break. For this, the members of Turing's group, led by him, familiarised themselves with the intricate characteristics of the Enigma cipher during the autumn of 1939, and soon became experts in the techniques used by the Polish. From there on, and using what in cryptography is called cribs, or when a fragment of plaintext may be linked to a fragment of encrypted text, he built an electric circuit to bypass the effect of the commutator, thus allowing him to ignore a septillion of its positions. This linked to the fact that they knew that

Enigma never coded a letter as itself, and aided by good machines also designed by them, which they called *bombe*, they were able to decrypt Enigma. Its importance was such that according to Harry Hinsley, without it the war would have ended in 1948, instead of 1945. To sum up, Turing once again used two heuristics to solve the combinatorial explosion, exactly as is currently performed in AI.

In 1943, Turing travelled to the USA for work-related reasons where he visited Bell Labs and met Shannon, the father of information theory. Being military secrets, they were unable to talk about their respective work and instead exchanged conjectures and opinions regarding the future of intelligent machines that were capable of learning. One of these meetings left Turing scandalised: Shannon wants to feed not just data to a Brain, but cultural things! He wants to play music to it!

That same year, 1943, three theoretical works were published on what is today known as Cybernetics.

In the first, Rosenblueth, Wiener and Bigelow suggested different forms and ways of endowing machines with purposes and goals, making them teleological and introducing the following three fundamental concepts:

1. Feedback as an organisational and control principle.
2. Teleological computation or with purposes and goals.
3. Information as pure form or abstraction, separated from the physical signal that transports them.

In the second, McCulloch and Pitts demonstrate how to model the nervous system and more. It is effectively on one hand, the initial stage of a theory of the brain that seeks solutions at the physical level, where structures and functions coincide. On the other hand, it serves to explain the problem of intentionality. Computation may be extensional, intensional and intentional or teleological as the most genuine aspect of human thought is linked to intentional computing, through understanding, and by the use of purposes and motivations to guide action. In this seminal work, McCulloch and Pitts apply symbolic logic to the problem of describing what neural networks are capable of. To do so, they demonstrated that all processes that may be described with a finite number of symbolic expressions such as basic arithmetic, classification, storage and recovery of finite datasets and the recursive application of rules of logic and so on can fit in networks, which they term formal neurons. These networks are a highly simplified representation of natural neural networks. For example, they are synchronous, that is to say, commutation occurs only during regular discrete intervals of time; that is, formal neurons are simply logical commutators. Nonetheless, they are important because they include any operation and process that may be described in terms of logic.

In the third, Craick, apart from proposing that machines use models and analogies to solve problems, succinctly established the idea on which rests the development of AI: That thought parallels reality through symbolism. Craick distinguished between three basic processes of human reasoning:

1. Translation of entities and relationships from the external medium, stimuli, to an internal representation in terms of words, numbers or other symbols such as, for example, images.
2. Deriving new symbols from older ones by means of processes of inference and any other cognitive processes in order to thus derive new internal representations.
3. (Re) translation of new symbols derived in terms of actions that work upon the external medium. Alternatively, the action may be replaced by acknowledgement at the reflexive level of the correspondence between these new inferred symbols and the events of the external medium to which they correspond, realising that a prediction has been fulfilled, assessed in this dynamic model of the medium. That is to say, to use the model of the medium, the knowledge base of each one, to predict and assess different alternatives that are converted into the most advantageous actions and update the model of the medium by learning. Craick made it very clear why he considered these to be the ideal requirements for designing an intelligent agent, when he stated: If the organism contains in its head a small-scale model of the external reality and of all its possible actions, it will be able to test different options, decide which is better, prepare its reaction against future situations that may emerge, use that which is learnt from past experiences and present and future situations, and in all situations, to react to imponderables in a satisfactory, efficient and risk-free manner.

Likewise and to a certain degree, Craick may be considered to have introduced the basics of the symbolic paradigm by interpreting human knowledge in terms of declarative and modular descriptions of high-level symbolic entities and of a set of inferential rules used to manipulate said restrictions.

Two years later, in 1945, Turing wrote an undated letter to Ashby where he said that in the course of his work on building the ACE computer, he had become more interested in producing models of brain function than in its practical applications. And in Proposed Electronic Calculator, he wrote that by trial and error, a machine could learn to play chess well.

Three years later in 1947, he gave a lecture at the London Mathematical Society where he mentioned computer intelligence, providing a surprising perspective of a new field. He described the human brain as a digital computing machine and discussed the perspective of machines acting intelligently, learning and winning over human opponents at chess. There and then, he established what a machine that could learn from experience would be, and to let the machine change its own instructions, endowing it with the mechanism to do so.

Halfway through 1947, Turing took a sabbatical year before he left the National Physical Laboratory (NPL). His report on the research conducted during this year was written in 1948 but remained unpublished until 1968, as his chief Charles Darwin did not consider it worthy of publication. Nevertheless, this report may be considered, in all certainty, the first manifesto on AI. Apart from other futurist ideas such as walking robots, Turing proposed new computation models, which he called unorganised machines or u-machines. These machines were of two types: One based on Boolean networks and another based on finite-state machines. Turing's inspiration for these two types of machines came from the workings of the human cerebral cortex and its ability to adapt itself. Likewise, in this report, Turing introduced a large number of ideas and concepts that would later be key to AI. Some noteworthy ideas among these are: A logic-based approach to problem-solving used widely today in knowledge-based systems; the concept of evolutive or genetic search, today called genetic algorithms; and a theory of the mind. This theory includes three of the key ingredients in the current theory of the mind. One is the suggestion that it is possible to say something useful regarding the functioning of the mind, which was a radical vision in those times of obscurantist behaviouralism. Two, according to Turing, the mind cannot be described relatively independently of what the brain does. And three, he advocated for a wide-ranging computational theory of the mind.

During the war, Turing experimented with “programs” for playing chess. Programs is placed within quotes due to the fact that, in the absence of computers, the machine’s behaviour was simulated manually, by pen and paper. In 1948, Turing and his friend Champernowne also built a flexible system of rules for playing chess that they named, somewhat unimaginatively, Turochamp.

In 1949, Hebb provided a simple rule of action to modify the intensities of the connections between neurons in neural networks of the McCulloch-Pitts type, such that they could learn.

One year later, in 1950, Shannon wrote a programme for playing chess. In this article, he stated that apart from performing numerical calculations, given their generality and versatility, new machines could also adapt themselves to symbolically work with elements representing words, propositions and other conceptual entities. This article also contains a lucid description of the problems of symbolic AI, namely, that brute force methods do not work and that learning was necessary, which his machine lacked.

That same year, Turing expanded his 1948 report for the NPL and published what was his most famous article, which began with the provocative question: Can machines think? Curiously, years ago in 1899, Father Leo Wiener, professor of languages at Harvard distributed among American readers, the poems of Rosenfeld in Yiddish. The first stanza of the first poem in the collection, titled In the Factory, apart from anticipating the question of the human use of human beings, also the title of a well-known book by Wiener, also poses the question of how machines might achieve thought. These are the verses:

> Oh, here in the stop the machines roar so wildly,
> That oft, unaware that I am, or have been,
> I sink and am lost in the terrible tumult;
> And void is my soul … I am but a machine.
> I work and I work, never ceasing;
> Create and create things from morning, will e’en;
> For what?-and for whom- Oh, I know not! Oh, ask not,
> Who ever has heard of a conscious machine?

Given the difficulty of unequivocally establishing the meaning of terms such as machine and thinking, a prior condition when responding directly to this question, Turing replaced the question with a Victorian English game called the Imitation Game, which was played at all social gatherings of the time. He described the game as follows: It is placed with three people, a man (A), a woman (B) and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A". We now ask the question: What will happen when a machine takes the part of A in the game. That is to say, Turing adopted a behavioural approach and proposed the following thesis: Although a machine may not be intelligent, its behaviour can be intelligent. In other words, if a machine were to win this game or if the interrogator is unable to distinguish between the machine and the other participant in the game, then the machine passes the test and consequently behaves intelligently. Conversely, and this is a frequently misinterpreted point, if the machine does not pass the test, then nothing is concluded. Later the test was called, in his honour, the Turing Test. That is to say, in the same way that Turing, when faced with the question of undecidability in 1936 when he was unable to formally define the concept of algorithm, invented his a-machines, he used something as well-established as the imitation game to define intelligence and its measurement.

One year later, in 1951, the aforementioned Minsky and his colleague Edmond, built the first neural network computer as part of their PhD. It was named SNARC and it used 3000 valves and an autopilot mechanism from a stripped B-24 bombardier to simulate a network of 40 neurons. As it happens, two of Minsky's PhD supervisors were sceptical whether this type of work could be considered mathematics. However, von Neumann who was also his supervisor predicted that: If it isn't yet, some day it will be.

That same year, Strachey built a program to play checkers, which included a heuristic search to reduce the combinatorial explosion that is generated by brute force methods. Meanwhile, Oetinger built the first program that explicitly incorporated learning.

Two years later, Shannon published an article, a continuation of that published in 1950 where he states: The problem of how the brain works, and how to design machines that simulate its activity, is surely one of the most important and difficult problems that are faced by science today. This article also poses a series of questions that constituted an authentic AI program: It included the following:

- Is it possible to organise machines in a hierarchy of levels, as the brain seems to be organised, and to enable the machine's level of learning to make gradual progress in this hierarchy?
- Can a computer be programmed such that eventually, 99% of the instructions that it uses are programmed by itself, instead of the current low percentage?
- Is it feasible to build a machine that can repair itself?

In 1955, Newell, Shaw and later Nobel laureate Simon developed the first AI language, IPL, which they used to build The Logic Theorist, which used heuristic search techniques to solve problems, and of which Simon declared: [We] invented a computer program capable of thinking non-numerically, and thereby solved the venerable mind-body problem. Subsequently the program was able to prove most of the theorems of Chapter Two of the *Principia Mathematica* by Russell and Whitehead. Russell apparently was delighted when Simon showed him that the program had achieved a shorter proof of a theorem than the one included in the *Principia*. Nonetheless, the editors of the *Journal of Symbolic Logic* rejected the article on this topic which included as authors Newell, Simon and The Logic Theorist. And thus we come to the summer of 1955, specifically 31 August, date on which the proposal for funding, signed by McCarthy, Minsky, Rochester and Shannon, was submitted with the title "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence" to the Rockefeller Foundation in the following terms: We propose that a 2-month, 10-men study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. 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. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. Next, they listed some aspects of the problems of artificial intelligence. Concretely the following seven aspects: Automatic Computers; How Can a Computer be Programmed to Use a Language; Neural Nets; Theory of the Size of a Calculation; Self-Improvement; Abstractions; and Randomness and Creativity. The total amount of funding sought was around $13,600 (McCarthy et al., 1995).

It is worth pointing out that firstly, the term AI appears for the first time in the title of the proposal, as until then this field of study was known as machinery intelligence, which was the term used by Turing. And secondly, its unbridled optimism which brought about a new age in this area.

Of this little-known era of AI, the most notable feature is the relevant and essential role that Turing played in it, to the point that it might be said, without exaggerating, that he and his machines as the basis of AI, his programs as practical performances and his test as the benchmark, can and should be considered the father of AI.

1.2.2 E2: Torrents of Optimism and Hope (1956–1968)

Keeping in mind the limitations of computers, both with regard to software and especially hardware, during that period, AI had a plethora of successes, albeit somewhat limited in scope. It must not be forgotten that, a little earlier, computers were considered machines that could only perform arithmetic calculations. It is thus understandable that the fact that they could perform more intelligent functions caused quite a stir. This made the intellectual establishment affirm that a machine could never perform X. These Xs included: Have a sense of humour, fall in love, taste strawberries with cream, what's interesting is that if it weren't for Turing's tendency to literally interpret everything, it might be said that he was being sarcastic with the detractors of AI. In response to these criticisms, AI researchers proceeded to demonstrate one X after another. For example, in 1958, McCarthy published a paper describing a hypothetical program that may be considered the first complete AI system. Similar to Logic Theorist, it was designed to search for solutions to problems, but unlike its predecessors, it used the general knowledge of the world. That same year, McCarthy defined LISP, after FORTRAN, the high-level programming language which became the most used for decades in AI.

One year later, in 1959, Gerlernter built a program, the Geometry Theorem Machine, which proved theorems using explicitly represented axioms. But given the resulting combinatorial explosion, he had to add an ad hoc heuristic. In 1961, Newell and Simon presented their General Problem Solver or GPS, designed to imitate protocols for solving human problems. With a few additional refinements, the GPS incorporated the Aristotelian approach, mentioned in his *Nicomachean Ethics*, of means-ends analysis; that is to say, a heuristic method that classified things according to their function and ranging from goals, required actions and means to fulfil them, to achieve the sought-after goal. It is true that this heuristic is unable to establish what is to be done when there are various actions for the same result, or when there is no action that enables the achievement of what is sought. Interestingly, Descartes's disciple, Arnault proposed in 1662, a correct formula for deciding the action to be taken in cases such as those mentioned above.

In 1960, Widrow and Hoff, and two years later, Widrow, improved upon Hebb's learning methods and called their neural networks ADAptive LInear NEtwork, ADALINE. In 1962, Rosenblatt presented his Perceptron, demonstrating its convergence theorem, with which he proved that the proposed learning algorithm was able to adjust the intensities of the perceptron connection to correspond to any input data, provided the correspondence was feasible.

From 1963 to 1968, Minsky's PhD students—Slagle, Bobrow, Evans and Raphael—developed various programs in the following areas and with the following names: Integration, SAINT, Algebra, STUDENT, Geometric analogy, ANALOGY, Semantic information recovery and SIR.

1.3 The Passage from Power to Knowledge and Learning (1969–1990)

In spite of the previous period's optimism, as represented by Simon's prophecy made in his Inaugural Lecture as President of the Operations Research Society of America in 1956. There and then, Simon declared that within the next ten years, that is, before 1967, computers would defeat the world chess champion, compose aesthetically pleasing original music and prove new mathematical theorems. All of it, following the touchstone of Turing and his test, to assess the intelligence of machines. As we know, when prophesying, Simon made a mistake that no experienced prophet would commit, namely, set a deadline for the fulfilment of the prophecy. Problems and difficulties soon began to emerge.

The first setback emerged with the ELIZA program developed in 1965 by Weizenbaum, through which, at least apparently, he could hold a serious conversation on any topic. The issue is that although it convinced some people, it lacked general knowledge and common sense, thus resulting in its failure.

The second setback was due to the unsolvability of many problems that were sought to be solved by AI. Most of the initial AI programs were based on the representation of the basic characteristics of a problem, and different steps were tested, until the winning combination that would produce an awaited solution was found. It worked initially due to the fact that the micro-worlds that were worked with had very few objects, that is to say, there was no combinatorial explosion. This made them think that all that was needed to advance from simple problems to complex ones was to install more powerful computers, which were faster and had more memory. This was false. It was the inability to deal with the combinatorial explosion, one of the main criticisms of the Lighthill report of 1973, that was used by the British government to withdraw its support for AI research, with the exception of two universities.

Finally, the third problem emerged as a result of the limitations inherent to the basic structures used to generate intelligent behaviour. For example, in 1969, Minsky and Papert demonstrated that although it was possible to make a perceptron learn anything that it could represent, its representational capacity was highly limited. A specific case was the impossibility of making a perceptron with two inputs discern whether the inputs were different. What was paradoxical about this case was that although the results of Minsky and Papert's work were not applicable to multi-layer neural networks, the funding for research into neural networks was reduced to practically nothing. Ironically, the new backpropagation learning algorithms used in multi-layer neural networks that were responsible for the spectacular resurgence of research into neural networks in the late 1980s were actually discovered for the first time in 1969, by Bryson and Ho.

1.3.1 E3: The Industrialisation of AI: Expert Systems and the Rehabilitation of Neural Networks (1969–1990)

The problem-solving strategy used in AI in the previous periods was based on general purpose search mechanisms in which steps of elementary reasoning and heuristics, to control the combinatorial explosion, were interwoven for complete and general solutions. This strategy was known as the power strategy and the procedures that used it were called weak methods, given that they scarcely used information, above all in the form of knowledge, if any. The result was that their performance left much to be desired in many complex domains. In order to achieve this improved performance, the best knowledge, when possible expert knowledge that existed in every domain of application was used. Thus passing from a strategy of power to a strategy of knowledge. The first system, based on expert knowledge, built with this approach was DENDRAL and its designers were Buchanan, Feigenbaum and the Nobel Laureate Lederberg. With DENDRAL, the complex and complicated problem of inferring a molecular structure from information in the form of data and news provided by a mass spectrometer, was undertaken. The system was fed the molecule's formula, for example, $C_6H_{13}NO_2$, and from there the mass spectrum provided information regarding the masses of the different molecule fragments that were produced when bombarded with an electron beam. For example, a mass spectrum with a peak of m = 15, would correspond to the mass of a methyl fragment, CH_3. The first version of DENDRAL generated all possible structures that corresponded to the formula; later it predicted the mass spectrum that would be observed in each case and compared them to the real spectrum. It soon became evident that this method was unworkable for medium-sized molecules and even less feasible for large ones. After consulting with expert analytical chemists, the creators of DENDRAL grasped that they worked by looking for well-known peak patterns in the spectrum that suggested common structures in the molecule. Thus, for example, to identify the sub-group of ketones, C = O, they used the following expert rule: If there are two peaks in x1 and x2 such that x1+x2 = M+28, M being the molecular mass, and x1–28 is a high peak, and x2–28 is a high peak, and at least one of the x1 and x2 is high, then there is a ketone group.

After ascertaining that the molecule contains a particular substructure, the number of possible candidates is considerably reduced, that is to say, the combinatorial explosion is reduced. All of this allowed the designers of DENDRAL to conclude that it had great potential, given that all the required theoretical information to solve this type of problem was correlated, from its general form in the component predicted by the spectrum, initial principles, with special effective forms, cooking recipes.

DENDRAL's claim to fame lies in its being the first intensive knowledge system that worked: Its knowledge was based on a considerable number of rules intended for special purposes. Subsequently built systems incorporate the scheme of clearly separating knowledge in the form of rules for the part corresponding to reasoning.

From there on, the DENDRAL strategy began to be used in other areas of human activity, according to the following heuristic, often used in invention: Now that we have this solution to a specific type of problem, what other problems can be solved with it?

Naturally, the first to do so were the builders of DENDRAL, and in the area of medicine. Buchanan and Feigenbaum, now accompanied by a doctor, Shortliffe, set to work and built MYCIN, an expert system for blood-borne infections. The system, which had 450 rules, made diagnoses that were as accurate as those of an expert. It had two main differences with DENDRAL. One, unlike DENDRAL, the rules for MYCIN did not have a theoretical model from which to make deductions. It was therefore necessary to obtain them after extensive consultation with experts, who had obtained them previously thanks to their direct professional experience in different cases. Two, these rules should reflect the uncertainty inherent in medical knowledge. This was achieved by including certainty factors in the system, which corresponded well to the way in which medical professionals weighed the evidence when making a diagnosis.

From this moment onwards, there was a proliferation of expert systems in all spheres of human activity. For example, in 1979, Duda and his colleagues built PROSPECTOR for mining prospections which was highly successful at recommending exploratory perforations at a certain site where significant deposits of molybdenum were found. Three years later, McDermott built the first commercial expert system R1, for the computer manufacturer Digital Equipment Corporation, DEC. This system was used to fulfil the orders for computers received by the company. In 1986, the use of R1 netted DEC 40 million dollars in savings. Today, the use of expert systems is commonplace and a matter of routine, such that the knowledge strategy was responsible for the paradigm of the same name. We shall look at the structure, functioning and mode of construction of these systems in Chap. 2.

In spite of the pause in the field of study on neural networks, after the article by Minsky and Papert, researchers such as Hopfield, Rumelhart and McClelland continued their indefatigable efforts to develop memory models based on neural networks. It was in the 1980s when the development of these networks received its strongest boost, when at least four different groups reinvented the backpropagation learning algorithm which, as we have mentioned, was discovered in 1960 by Bryson and Ho. Since neural networks will be dealt with in more detail in the following chapter, we shall speak no more of them here.

1.4 The Elements of an AI Theory

AI may be viewed as a science, in which case, the object of study is to build an intelligence superior to human intelligence outside its natural habitat, the human brain. From this perspective, the important questions posed are: What would be its most suitable platform? Is a general AI possible? What would be the constituting elements of such a theory?

In this regard, in 2013 the authors showed how information along with energy and matter is one of the three basic building blocks of the universe and therefore, interchangeable with the other two in the same way the matter and energy are interchangeable according to the formula: $E = mc^2$. From there, they propose two essential elements for the establishment of a theory of AI:

- Holons are entities that have the capacity for autonomous behaviour, that is to say, behave as a whole, but not self-reliantly, thus behaving as part of a bigger place. Insert paragraph corresponding to Section B of our article A new approach … published in Foundations of Science.
- Informons, etymologically speaking, means to 'give shape' in Latin inform, and 'part of' in Classical Greek.
  The Informons and holons can be combined as follows:
  — holon x holon ——> a more efficient holon.
  — holon x Informon ——> a more elaborate and a better quality informon. That is to say, if it were a date, it will become news, and if it were news, it will become knowledge.
  — informon x holon ——> a more efficient holon.
  — informon x informon ——> a more elaborate informon.

This, united with the postulates of complementarity, computation and satisfiability, complete the theory.

If, in line with Couffignal (Couffignal, 1969), we define theory as a classification whose classes contain definitions that explain each other. Theories are established in one of the following two ways:

A. Empirically. In this case, observing and identifying the characteristics of certain beings and subsequently creating increasingly abstract or general classes within which these beings are accommodated. These theories are called empirical and they are akin to the classifications made by naturalists and are based on elementary mental models created by natural beings, that is, not created by humans.

B. Axiomatically. Here, sub-classes are built from the abstract definition of the class, adding arbitrarily selected properties to the definition of this class, with the sole proviso that there are no contradictions. These theories are called axiomatic theories and are similar to mathematical or logical theories and are based on mental, dialectal models of beings invented by the human being.

The attributes to be added to the comprehension of a class of beings in order to achieve the understanding of a sub-class, constitute what is called the technology of said sub-class in relation to the considered class. Normally, only one classification is considered and the technology of a sub-class is defined in relation to the most abstract class possible. In this way, the term technology is used in an absolute sense.

When building a theory, the technological properties of a being are not considered as belonging to a certain class, as these properties are not deduced from the definition of the considered class. Consequently, it may be affirmed that a deductive reasoning cannot discover the technological properties that have been ignored when developing the theory. For example, the theory of the resistance of materials, which is of use when calculating the construction of a bridge, assumes that the terrain on which the bridge is supported is perfectly rigid and infinitely resistant, ignoring the physico-chemical properties of the land itself. Therefore, calculating a specific bridge by methods dealing with the resistance of materials cannot include and account for these properties. Therefore, it may be affirmed that the deductive theories that are found in almost all scientific theories, represent the observed phenomena very imperfectly. Hence, in order to safeguard the framework of reasoning that constitutes a theory, it is necessary either to disregard concepts that cannot be explained by the theory, as in the case of aether drag by matter, or to attribute to the beings under observation properties that they do not possess, as in the case of the ether until it was banished by Einstein. A theory is, therefore, a mental edifice or construction established on the basis of observed phenomena, but it does not constitute an absolutely faithful image of said phenomena.

The different classes into which natural beings are classified constitute new entities to which a symbol is attributed. In an axiomatic theory, the classes are the result of deductive reasoning, and a symbol is also attributed to them. When dealing with reasonings whose hypotheses are symbols or mental models of natural beings, it is more comfortable to use symbols that refer to the hypotheses. In this way, when these symbols are used, they evoke patterns considered to be the same as those of natural beings. These patterns constitute a reification of the definition of the symbols. For example, once we verify that all gases, under normal conditions, approximately verify the equation $PV = RT$, where P, V and T are respectively the pressure, volume and temperature of a gas, and R a universal constant, the physical entity that exactly verifies this law is called a perfect gas. The term gas evokes the pattern corresponding to an entity that exists in reality, and constitutes a reification of the formula $PV = RT$.

A reification thus consists of adding an invented technology to a theoretical result. However, the being of this invented form may not exist physically. Therefore, it is necessary to distinguish between physical or real existence, which is that of a being whose definable attributes are observable by known physical means, and fictional existence which may be attributed to a being whose definable attributes do not contradict known and accepted laws that are applied to natural beings. Thus, a reification of a theory passes from fictional or mental existence to physical existence by means of an experimental check. Usually, sciences use modifications that only have fictional existence. This is the case with entities or concepts defined as causes of physical phenomena. For example, heat causes variations in temperature; forces are responsible for variations in speed; light, optical phenomena and so on. Proof of this is, for example, that in 1955, Denis Gabo presented an optical theory in which the substance commonly called light played no role. These causes considered by the natural sciences are defined by attributes that are common to a set of beings with physical existence who are capable of action, for example, in the case of heat on a thermometer. This action is expressed by stating that, with the exception of a change of state, the application of heat to a body implies an increase in its temperature. Nevertheless, it must be remembered that thermodynamically, the passage of heat from a body to another body is only possible if there is a difference in the temperature of both bodies. What's more, heat is usually considered an energy in transit due to the thermal gradient. In any case, the implication appears as the dialectic expression of the relationship between the cause and effect.

We must not forget that for praxeological purposes and practical implementation, a theory by itself, without the suitable technology for a correct reification, may have undesired or even dramatic results. A theory offers a set or a series of models that are deduced from each other, and all the eliminated attributes to build the models, which belong to the predicate fields of the beings in question. For a given being, the attributes to be added to its model to reconstruct the original, is the technology of this model. A dramatic example that occurred in France many years ago, demonstrates the importance of the technology. It was necessary to build an elevator in a mine to transport men and materials from its interior. It was done normally, as shown in the figure, with a shaft and a sleeve on which the elevator drum was located: First phase of reasoning. Then the diameter of the shaft was calculated, second phase of reasoning. Next came the third phase, building the original on the basis of this data. Finally, it was checked, after various tests, that the elevator could bear all the weight that was placed in it and much more. Thus it was concluded that the calculations were correct and the elevator would work without problems, that is to say, the scientific study of the elevator was perfect. But then a tragedy occurred. After functioning normally for a few years, the elevator broke and 11 miners fell down to a depth of 250 m and died. What happened? To put it simply, a technological point was lacking. There should never be an acute incoming angle in a mechanical construction. Experience teaches and technology corroborates that incoming angles pose a breakage risk and as shown in the Fig. 1.1, should be replaced with carets. The engineer responsible for the elevator project was perfectly aware of rational mechanics, elasticity and the resistance of materials, but he either ignored or forgot this technological aspect that there should never be an incoming angle.

This is also a good example of the rule that when considering an abstract model in relation to the original, one must think and take into consideration all the technology that was not included. This notwithstanding the fact that often, when dealing with the abstraction, one may throw the baby out with the bathwater.

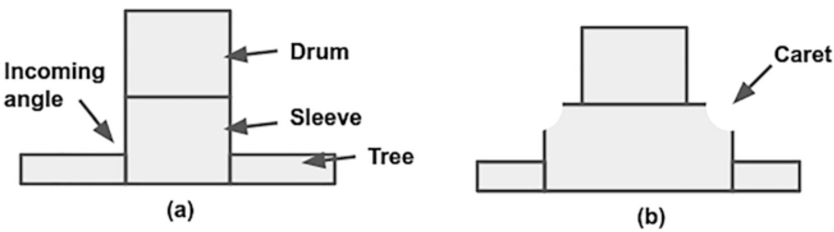

Fig. 1.1 Importance of Technology

1.5 Fundamental Framework for AI

As mentioned in Sect. 1.2, there are three relevant approaches to working with AI currently:

- The symbolic approach, which is the closest to traditional computer science. It is based on formal logic, concretely on deductive inference, and it is easy to understand, as it often presents a declarative form as in the case of expert systems. Chapter 2 shall look at its architecture and mode of functioning, as well as its building method.
- Neural networks, hereinafter, NN when singular, and NNs when plural. This approach is somewhat more difficult to understand intuitively, and is inspired by the brain's physico-chemical functioning. It uses several layers of artificial neurons that learn and compares them to the desired result and altering the weight of the connections between the neurons, such that the most ideal connections are strengthened. This approach uses inductive inference. Likewise, Chap. 2 shall delve into the structure, learning and method of building NNs.
- The Bayesian approach. This approach uses plausible or abductive inference: Given the evidence, what is the probability that the hypothesis or conjecture is more true. For which, in Chap. 7, we shall examine a real-life case of its functioning applied to the hypothesis of the Technological Singularity.

In accordance with the postulates of the theory proposed in Sect. 1.3, these approaches can be integrated with each other and with evolutionary systems. For example, NNs can learn to make inferences, generalise and to make analogies. As a matter of fact, increasingly more AI systems implicitly or explicitly use several or all of these approaches along with other algorithmic or heuristic procedures. What is undebatable is that the best approach is the one that is ideal to solving the ongoing problem.

References

Anguera de Sojo, A., Ares, J., Lara, J. A., Lizcano, D., Martínez, M. A., & Pazos, J. (2013). Turing and the Serendipitous Discovery of the Modern Computer. *Foundations of Science, 18*(3), 545–557.

Anguera, A., Lara, J. A., Lizcano, D., Martínez, M. A., Pazos, J., & De la Peña, F. D. (2020). Turing: The Great Unknown. *Foundations of Science, 25*, 1203–1225.

Borrajo, D., Juristo, N., Martínez, V., & Pazos, J. (1993). *Inteligencia Artificial. Métodos y Técnicas*. Editorial Centro de Estudios Ramón Areces.

Couffignal, L. (1969). *La Cybernetique*. Presses Universitaires de France.

Guralnik, D. B. (1969). *Webster's New World Dictionary of the American Language Enlarged from the Concise Edition with Student Handbook*. The Southwestern Company.

Hodges, A. (1985). *Alan Turing: The Enigma of Intelligence*. Unwin PaperBacks.

Lara, J. A., Pazos, J., de Sojo, A. Anguera, & Aljawarneh, S. (2021). The Paternity of the Modern Computer. *Foundations of Science*, On-line. https://link.springer.com/article/10.1007%2Fs10699-021-09797-y#citeas

McCarthy, J. M., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955, August 31). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.

Rich, E., & Knight, K. (1991). *Artificial Intelligence* (2nd ed.). McGraw-Hill, Inc..

Russell, S., & Norvig, P. (2005). *Artificial Intelligence a Modern Approach*. Prentice Hall Inc.

Teuscher, C. (Ed.). (2004). *Alan Turing: Life and Legacy of a Great Thinker*. Springer-Verlaine.

CHAPTER 2

AI: Methods and Techniques. Knowledge-Based Systems

David Lizcano Casas and Juan Pazos Sierra

Abstract The acquisition, management, processing and improvement of knowledge, as well as the effective and efficient treatment of huge amounts of data, big data using learning techniques, have become a very important function of information systems applications within organizations for solving problems with Artificial Intelligence.

Rapid technological progress has been made both in knowledge-based systems, including expert systems, modes reasoning and knowledge acquisition and management, as in Artificial Neural Networks, and its construction methodologies, which is essential in order to deal , respectively, with knowledge assets and huge amounts of data in organizations currently. This chapter relates knowledge-based systems and artificial neural networks for the management and increase of organizational knowledge and specifies different rules, methods, techniques and strategies that can be adopted by companies for the successful implementation within Artificial Intelligence applicability scenario.

D. L. Casas (✉) • J. P. Sierra
Universidad a Distancia de Madrid (UDIMA), Madrid, Spain
e-mail: david.lizcano@udima.es; juan.pazos@udima.es

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_2

Keywords Artificial Intelligence • Problem-solving • Knowledge-based systems • Expert systems • Production systems • Modes of reasoning • Artificial neural networks • Organisations • Strategy

2.1 Introduction

Undoubtedly, problem-solving is the main mission of Artificial Intelligence (AI), a term that is derived from the Greek language, and which may be described as stating that an entity, whether natural or artificial, has a problem when it wants something and does not immediately know the action or series of actions to be performed in order to achieve it. The desired goal may be something tangible, such as to eat an apple, or abstract, as in the proof of a theorem. That is to say, it may be a physical object or a set of symbols. The actions involved in obtaining the desired objects include physical actions, perceptual activities and other purely mental activities, such as judging uniqueness and memories.

To have an idea of what exactly constitutes a problem is also to specify the conditions for its existence and its components. In the simplest of cases, these conditions are as follows:

1. There must be at least one individual, man or machine; it is initially irrelevant which, to whom the problem may be attributed, within a frame of reference or environment. This framework is defined by uncontrollable variables.
2. The individual must have at least two alternative courses of action, that is they must be able to make a choice of behaviour. A course of action is defined by one or more qualitative or quantitative values of the controllable variables.
3. There must be at least two possible outcomes of the selection, one of which must be more acceptable than the other. In other words, there must be, at the very least, a desired outcome, that is an objective or goal.
4. The selection of any of the solutions must have a different impact on the objectives of the system, that is there is an efficiency and effectiveness associated with each solution that must obviously be different. The individual is then said to have a problem only if they do not know what the best course of action is and wish to know it. Moreover, they must have doubts regarding the solution. To sum up, it may be stated that an entity has a problem, if they want something, they have alternative forms of

seeking that something, but with different levels of efficiency for achieving it, and doubt regarding the best course of action. Of course, problem situations may be much more complex than the one described above, for example, the problem may involve a group, rather than an individual. The environment, or frame of reference of the group, changes dynamically in ways that affect the effectiveness of courses of action, or outcome values. There may be many, albeit finite, alternative courses of action and the number of goals may also be very large. Additionally, these goals may not necessarily be consistent. Alternatives may be implemented by another unrelated group, but not independently of the problem. The effects of the decision taken by the group affect other groups, whose reaction may be favourable or unfavourable.

In other words, a problem is not fully defined or unworthy of formal consideration unless it meets the following conditions:

(a) Expressed in terminology clearly understood by its potential solver.
(b) The form and notation for the solution must be agreed upon.
(c) Have a way to identify the relevant data on which a solution may be based.
(d) To establish conventions regarding how to measure the validity and acceptability of proposed solutions.

According to an old aphorism, a well-framed problem is a problem half solved. That is evident. What is not so obvious, however, is how to frame a problem well. In its broadest formulation, any problem **P** may be stated by the following triple:

$$P = < S, O, G >$$

where **S** is the set of states or expressions representing objects, as well as assumptions, definitions, axioms, facts and the like. Often, we start from a privileged state: The initial state represented by Si. **O**, on the other hand, represents the set of operations or actions that can be carried out on these states and has a twofold mission. In action-oriented problems, it is the set of actions or transformations that can be performed on **S.** In conclusion-drawing problems, the operations are given the particular name of inference rules. Whereas in proof problems, the operations are called rules of inference.

It is necessary to distinguish between two types of operations. The first, called destructive operations, produces new expressions at the cost of old ones. The second, called constructive operations, produces new expressions by increasing the set of existing operations, without destroying old ones. Operations in logical and algebraic problems belong to this second type.

Finally, **G** is the terminal expression or goal to be achieved. There are two paths or approaches to obtaining systems or artefacts that are capable of solving problems:

One, the symbolic path, which seeks to simulate the workings of the brain by considering it an information processor. Consequently, it seeks to create systems whose behaviour is such that if carried out by a human being it would be considered intelligent. What is of interest here is using the capabilities of computers, that is their computational speed, memory capacity and reliability, to the maximum, without worrying about whether they use the same methods as biological systems, similar to how planes fly differently from birds yet perform better in their goal of moving and transporting people and objects through the air. Within this approach there are two complementary points:

(a) The heuristic, or power paradigm, which involves intensive computation. Thus, it uses the power of computers to perform computations and search processes, allowing complex and dynamic problems to be modelled.
(b) The epistemological, or knowledge paradigm, which considers AI the science of knowledge, that is how to obtain, represent and use general knowledge and, especially, common-sense knowledge, which is fundamental for the resolution of problems that arise in all types of activities that require intelligence to perform them. Knowledge-based systems are the most conspicuous manifestation of this approach, and will be discussed in greater detail below.

Two, sub-symbolic or emulation systems, also known as connectionist systems, which imitate, as far as possible, the structure and functioning of biological systems for behaviour that is considered intelligent. Its advantage is that the researcher is required to specify a model in full detail and it is also possible to derive consequences by simply running the model. This emulation approach is called sub-symbolic because knowledge is encoded in the connections of neurons and not explicitly given. Given the successes, some of them spectacular, which have been achieved recently using this approach in the form of artificial neural networks (ANNs), it will also be considered in more detail below.

2.2 Knowledge-Based Systems: Expert Systems

2.2.1 Production Systems

Production systems are one of those happy events that sometimes occur in science and technology: A well-established and effective formalism that is awaiting practical use.

It has a long and varied history. Formal research into the knowledge representation began in India with theoretical treatises on Shastric Sanskrit grammar, circa 1000 BCE. Indeed, the grammatical theory of this language proposed not just the formal vocabulary and syntax of a general-purpose language, but also included an analysis of its semantics. This made it the first example of systematic knowledge representation in a specific area, for the purposes of facilitating inference.

In the West, the first example was the use of definitions of ancient Greek mathematical terms. More recently, their use in symbolic logic began in the form of production rules, with Post after whom they are named (Post, 1947). They were also shown as “Markov algorithms”. And, to come full circle, Chomsky used them in linguistics, where they are called “rules of rewriting”. Finally, they were taken up again in the late 1960s to “programme” a wide variety of experiments.

The success of this form of representing knowledge is due to the fact that its form: “IF … THEN”, is at the root of all behaviour, whether electronic or human. Consider this passage from Shakespeare’s Richard III (Shakespeare, 1989):

(a) Buckingham: “Go, gentle Catesby
And as it were far off, sound thou Lord Hastings
How he dote stand affected to our purpose;…
…
If thou dost find him tractable to us,
Entourage him, and tell him all our reasons:
If he be leaden, icy, could, unwilling,
Be thou so too, and so break off the talk,
And give us notice of his inclination;
…

Now, my lord, what shall we do, If we perceive
Lord Hastings will not yield to our plots?
(b) Gloucester: Chop off his head;".

As may be seen, Shakespeare omits the word "then" so as not to spoil the rhythm of the story.

On the other hand, many times, as in the narrative poem by Henry Wadsworth Longfellow (1807–1882): "The Midnight Ride of Paul Revere" (Longfellow, 1861), the "IF … THEN" pair proves to be the most efficient at transmitting information. In April 1775, the military reserve forces at Lexington and Concord, Massachusetts, were about to fall to the British who were advancing from Boston. The American commander Paul Revere needed to know whether the British were coming by sea or by land so that he could properly deploy his troops. He told his informant:

If the British march
By land or sea from the town to-night,
Hang a lantern aloft in the belfry arch
Of the North Church tower as a signal light;
One if by land, and two if by sea;
And I on the opposite shore will be,
Ready to ride and spread the alarm
Through every Middlesex village and fármacos,
For the country folk to be up and to arm.

Finally, there is the more than famous poem "IF" by Rudyard Kipling, as an example of the use of the rule scheme "IF … THEN" (Kipling, 1992):

If you can keep your head when about you
Are losing theirs and blaming it on you
If you can trust yourself when all men doubt you,
But make allowance for their doubting top;
If you can wait and not be tired by waiting,
Or being lied about, don't deal in lies,
Or being hated, don't give way to hating,
And yet don't look to good nor walk too wise:
If you can fill the unforgiving minute
With sixty seconds worth of distance run,
Yours is the Earth and everything that's in it,
And—which is more—you'll be a Man, my son!

2.2.2 Structure and Components of Rule-Based Systems

As shown above, the productions may be represented by the following scheme:

Condition→Action or, more generally,
Conditions→Actions, this is purely logical.

However, as shown in the examples above, when speaking or writing it is customary to use the If condition(s) Then action(s) scheme. In this case, the productions are called production rules or simply rules.

A simple but widely used example of a rule would be: IF the problem is solved THEN stop.

It is possible to establish two well-defined classes of productions: One general and another special, called a perception test. The condition of a production belonging to the general class is some kind of test usually on the presence or absence of a particular type of symbol in the content of the working memory. The test is either satisfactory or unsatisfactory. Provided that the test in the condition part of a general production is satisfied, only then can the action part be executed. Such an action may be a motor action, the execution of a perceptual test, a retrieval of information stored in a memory, or the action of changing the fact base content.

The scheme of the aforementioned productions may be generalised in such a way that the condition part of a production may involve a group of several elementary tests, and, in turn, the action part may contain a sequence of elementary actions. For example, let the following rule apply:

IF Card = verified and Date = not expired and PIN = correct and Attempts < 4 and Balance = sufficient and Limit < €3000 THEN Payment authorised and Show remaining account balance.

Almost all sequential programs are dependent on data, initial conditions, parameters, dynamic interactive responses from users or digitised representations of signals received from instruments, and results of previous computations. In these programs, control flow and data usage are especially rigidly specified by the program code. Their specific shortcoming is their sequential nature, as branching only takes place at points and paths that are explicitly provided for in the program code. While this structure is well suited for some types of computations, it is ill-suited for others, such as simulating human responses to a complex, rapidly changing and unfamiliar environment. In these situations, bifurcation may be the rule rather than the exception. That is to say, the program will examine the state of the world at each step and react appropriately. When no major new stimulus is perceived, the current context, history or last state may be the controlling factor that makes the program move one step further sequentially. However, if a major new stimulus is detected, the next step of the program will depend on the properties of that stimulus, as well as the content. In situations where new stimuli are continuously emerging, either from the external environment or from conditions generated internally by the program itself, direct sequential coding may be insufficient. In these cases, the program may rather be seen as a loosely organised collection of pattern-driven modules that are responsible for both detecting and responding appropriately to different situations. This is where rule-based systems come in (Borrajo, 1993).

Its basic components are (Gómez et al., 1997):

> A knowledge base, containing: Rules, that is an internal representation of the productions, as well as of concepts and static relations between them. And, meta-rules, to facilitate: The most efficient selection possible of the normal rules for execution. The use of ad hoc heuristics.

The acquisition of new rules. An inference mechanism, called an inference engine, which selects the applicable rules at each point in the running process and executes them in order to obtain a conclusion, that is to perform some kind of knowledge processing or interpretation.

Additionally, such systems contain a fact base or working memory or active memory that accumulates a set of established facts that are used to determine which rules can be applied by means of the inference mechanism. This set of facts may exist as part of the initial conditions of the system, as well as a result of the inference processes, and the different conclusions that can be drawn from this process will be added to it. Finally, new facts may come from sensors and databases outside the system. Changes to the system's status occur in this memory, such that it always represents the current status of the system. This is why working memory is responsible for interacting with the outside world by facilitating the input of non-inferential information. It is also the permanent focus of the system's rules. The fact base and working memory are generally considered to be the same thing. Therefore, we speak of a permanent fact base when it represents the declarative structure of the system, and a temporary fact base when it only includes declarative knowledge regarding the problem being

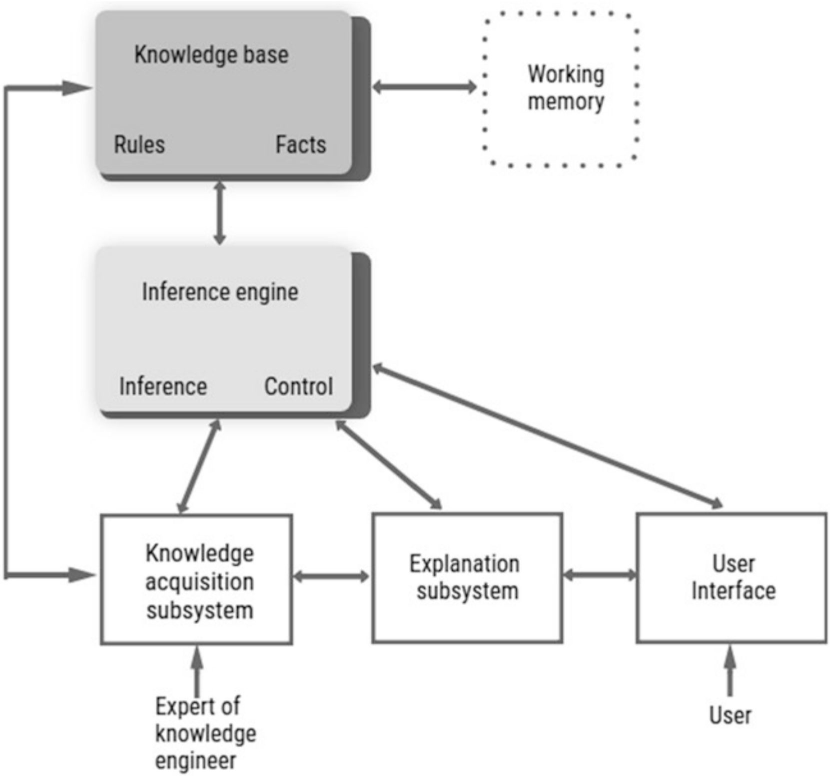

Fig. 2.1 Rule-based system's general outline

solved. However, at other times, for convenience or for any other reason, they are considered differentiated or even different entities. This simplifies the understanding of both structures and the problems arising from the relationship of the knowledge-based system with the outside world.

Figure 2.1 shows the general outline of a rule-based systems (Prerau, 1990).

The stages of action of the inference engine are as follows:

1. Decision or rule selection phase

   Selection of the rule to be executed, which first involves partitioning the Facts' Base, FB, and Rules' Base, RB, into different subsets. This constitutes the "restriction" stage. The goal of this stage is intended to accelerate the next stage. Next is the matching stage, which consists of selecting the set of BH-compatible rules that are candidates for activation.

There are several matching techniques, of which the temporal redundancy technique, RETE, is one of the most important. And finally, there is the "conflict resolution" stage. Since, as a result of the previous step, there are almost always more rules than processors that can execute them, it is necessary to choose which of these rule(s) will be activated. At this stage, different heuristics may also be used for greater efficiency. The following are among the most commonly used: Choose the first rule that matches the context; select the rule with the highest priority, the priority being set by the designer in consultation with the expert; choose the most specific rule, that is the one with the most detailed condition, or the strongest requirements; or choose the rule concerning the item most recently added to the working memory.

2. Action stage

It consists of effectively applying the selected rule(s). Generally, this phase is limited to introducing new facts into the BH and deleting or modifying others. However, the use of classical procedures is allowed both in the action part and in the rule condition in complex knowledge-based systems. This cycle is repeated until the target event or action is added to the BH, or when there are no rules that may be applied.

Finally, and although it does not strictly belong to the system, in order to enable communication with the environment, the system is equipped with a module for interacting with the outside world.

2.2.3 Modes of Reasoning

In short, the existence of rule systems and human modes of expression are compatible with and suitable for computers to handle the knowledge they have about any task at hand. This is how knowledge-based systems, hereafter referred to as KBS, came into being. And if this knowledge also contains expert knowledge of a specific area, these systems are then called expert systems, hereafter referred to as ES, and their design and construction constitutes what is known as Knowledge Engineering, hereafter referred to as KE (Gómez et al., 1997).

The term that differentiates between these and other intelligent systems is knowledge, which in turn is a specific type of the more general term, information. First of all, as the authors have pointed out and proved, this term refers to one of three basic concepts that make up the ontology of the world, which are interlinked with each other. The other two are matter and energy, which, as is well known, are equivalent in that they can pass from one to the other, as shown in the famous equation known as the fate of mankind: $E=mc^2$. Similarly, it is possible to pass from information to energy and vice versa (Barreiro et al., 2019).

However, the term information encompasses, at a first level, both the medium of information: Air, water, cables, paper, electronic devices and so on, and the signs and signals with which it can be expressed. On a second level, there are data, which represent the values that a variable can take; news, which, on many occasions, committing an obvious synecdoche is itself called information; and knowledge. An example will make this clear. Suppose a health worker takes a patient's temperature with a thermometer and it is 42 °C. This is a data point. It is well known that any human temperature above 38 °C means danger for the person involved. This makes it possible to take appropriate measures to at least alleviate the consequences of this imbalance and, if possible, to eliminate it, for example by applying cold compresses to the forehead and prescribing an antipyretic. This implies knowledge. That is, data has to do with the syntactic aspect of information, news with its semantic aspect and knowledge with its pragmatic aspect. However, it is not always easy to distinguish between these three aspects. Finally, there is a third level where wisdom, purposes and so on are to be found, which, at least for the time being, are beyond the reach of the KBS.

Now if we focus on knowledge; they may belong to the following three types: Explicit, which are already available to anyone who seeks access to them; implicit, which are often presented in languages that can only be understood by the initiated, for example, mathematical formulas or musical scores; and tacit, which are possessed and used by real experts in a domain on a daily basis, but which they are unable or unwilling to make explicit or even implicit.

In this regard, some paradigmatic examples of implicit and tacit knowledge and the difficulty of making them explicit are given below. The first concerns the theory of general relativity and its most important originator, Einstein. As is well known in Astronomy and Physics, the equations of this theory assumed the existence of black holes, gravitational lenses and so on.

Well, not only did Einstein fail to see them, he also tried to prove that they were impossible. And this happens relatively often, to the extent that it is now commonplace to say that many theories that are implicitly expressed with mathematical formulae are more intelligent than their own authors.

As far as tacit knowledge is concerned, the most striking case is undoubtedly that of Srinivasa Aiyangar Ramanujan (1887–1920), a mathematician, who achieved a numerous important and highly unconventional results, some of them not only illuminating but also fruitful (Kanigel, 1991). He attributed these results, the fruit of his intuition and insight, to his family deity, Mahalakshmi of Namakkal, from whom he said he received visions of scrolls with complex mathematical content that unfolded before his eyes. For him, an equation or formula was meaningless if it did not represent God's thought. In fact, he was a great creator of formulas, as recorded in his notebooks, many of them unproven and others that were proved by other mathematicians, especially Hardy. In short, Ramanujan knew, that is tacitly, many theorems, but not their proof which involved implicit and explicit knowledge (Ramanujan, 2012).

2.2.4 Knowledge Acquisition

This is perhaps the most important, difficult and costly process involved in developing an ES, and constitutes a true bottleneck for ES development. Certainly, there are many and varied sources of both explicit and implicit human knowledge. However, these sources lack the fundamental knowledge of ES, the tacit knowledge that is only in the heads of experts. The process of extracting knowledge from an expert, known as knowledge education, is highly problematic not just because of its inherent difficulty, but also because the expert may be reluctant to give up their expertise, mainly because it endangers their status, to the point of making them expendable. Therefore, the first thing to do to overcome this non-technical hurdle is to reach an agreement with the expert.

Next, there are various techniques to help the knowledge engineer perform their task of education. Some are manual, such as interviews, questionnaires, protocol analysis or the observation of the expert themselves in the course of their normal work. Others are semi-automatic, such as grating or psychological scaling. And the third are automatic, as in the case of the W algorithm to describe the characteristics of a concept by induction, or by using decision trees.

When one wants or needs to extrapolate knowledge from a group of experts, the following techniques may be used: Brainstorming, nominal group technique and the Delphi method.

All acquired knowledge, compiled if explicit, made explicit if implicit, and extrapolated if tacit, is included in the ES, either in the knowledge base or in the fact base, to be exploited by the inference engine, taking into consideration strategies, priorities, risks, utilities and probabilities, in order to obtain an optimal solution.

Apart from the three essential elements of an ES, these systems are provided with ad hoc input-output interfaces to communicate with users, physical devices, or with other ES, databases or other information systems.

Last but not the least, ES are lately implemented using ontologies as the fundamental support of the knowledge base and problem-solving methods as inference engines to make the ES as independent as possible so they may be reused to build other systems.

Obviously, there are methods to design, develop and implement these systems. Some of the most widely used are CommonKADS, IDEAL and KLICyMIKE.

2.2.5 The Knowledge-Network of Digital; Management of Organisational Knowledge

KBSs have a direct impact on the management of knowledge within organisations. A KBS aims to extract a piece of the organisation's intellectual capital, and capture it in an information system. Thus, the development of a KBS results in the creation of a tangible knowledge asset which can be distributed and leveraged within the company. The knowledge targeted by a KBS can be either tangible (such as manuals and documents) or intangible (e.g., human expertise) in nature. While a conventional information system can store the information in a document or a manual, the additional power of a KBS lies in its ability to also store how the document or manual is interpreted (used) by experts/users (i.e., capture the intangible knowledge associated with the use of the tangible knowledge asset).

2.3 Artificial Neural Networks

2.3.1 Introduction

One of the oldest problems in experimental science is to find functions that fit or explain data observed from natural phenomena. The main advantage of these functions is the possibility of predicting the future behaviour of the natural system and to control its outputs by applying appropriate inputs. Practical examples of this are weather forecasting, classification of tumour patterns or forms in medical diagnosis or predicting stock market behaviour. The difficulty is that observed data tend to be accompanied by noise, and the exact mechanisms that generate it are usually unknown. Occasionally, it may be possible to establish an exact mathematical model that fully explains the process from which the observed data originate. However, it is impossible to provide the details of this process on many more occasions. In this case, the objective will be to estimate the underlying model that generates the observed data. In these cases, machine learning techniques allow these models to be established using example data or past experience. This permits the identification of certain patterns or regularities in the data, and the construction of good approaches to the problem. These techniques, which may be called semi-parametric estimation, include ANNs.

2.3.2 Biological Foundations

The nervous system (NS) is considered (Ramón y Cajal, 1960) a unique structure, the most complex one in the universe. Its main function is, along with the endocrine system, to ensure control and communication within the organism. The nervous system receives approximately 10^{10} bits of information every second, of which only 10^2 bits are consciously admitted, 99% of the information received being ignored. The information that reaches the nervous system is encoded and then conveyed or transmitted and finally incorporated into a signal that determines the production of an intellectual, motor or neurosecretory work as a system response. In this last instance, the active and personal integration of the organism is achieved.

The nervous system only has two specific types of cells: Neurons and glial cells. Until recently, neurons were thought to be the basic functional element of the nervous system. However, neurons never actually appear in isolation; they always form functional structures through their extensions and in association with glial cells. A typical neuron in the nervous system has three well-defined parts: The soma, which is the cell body and the integrating area for the information that reaches the neuron; the axon, which originates in the conical area of the soma, and its terminal end produces several branches, that is it constitutes the information transmission area; and the dendritic tree, which also originates in the soma, is highly branched and constitutes the information receiving area.

As is well known, the program to construct the nervous system is encoded in the genes, but this coding is not sufficient to determine all possible situations and connections of the elements of the nervous system. It is through learning by repetition of responses to deterministic stimuli that renders the circuits of the nervous system specific, as the physical substrate of its acquisitions, altering existing synapses and making previously diffuse and overlapping projections precise, point by point. In this way, the so-called metacircuits and metastructures, as Barbizet called them in 1961, are formed, which are the supports of lived experiences, resulting in a new order of neuro-glial organisation.

Each neuron receives information from various receptors and/or neurons, the convergence principle; and integrates and distributes it by means of the axon with its branches to a series of neurons and/or effectors, the divergence principle. The functional structure of the neuron is designed so that it basically fulfils three partial missions with the information that reaches it, namely: It integrates it into its own activation code, it transmits it in frequency-coded form through its axon and, finally, it transmits the impulses to its target elements at its endings.

A fundamental element of the nervous system is the synapses which constitute the functional contact between two neurons. Each synapse consists of a presynaptic element, which belongs to the neuron sending the impulse, and a postsynaptic element belonging to the cell receiving the impulse, as well as an intersynaptic space between the two elements. There are basically two types of synapses in the animal world: Chemical and electrical. Most of them are chemical in nature, whereby a neuron secretes specific chemicals, generically called neurotransmitters, hereafter referred to as NTs, into the intersynaptic space, which will act on the receptors. These synapses have a very important characteristic that makes them ideal for transmitting signals in the greater part of the nervous system; they always transmit these signals in only one direction, from the presynaptic neuron to the postsynaptic neuron, which allows, among other things, signals to be sent to specific areas or points in the nervous system. Electrical synapses, on the other hand, are characterised by the fact that they are direct conduits for the passage of the electrical impulse from one neuron to the next, thus allowing the impulse to pass in both directions. These synapses are made of tubular protein structures called gap junctions that allow ions to pass freely from one cell to the next.

From what was just said, the synapse is deemed highly advantageous for controlling signal transmission, as it establishes its direction; it may facilitate, inhibit, weaken, the transmission of the signal, blocking, increasing, diversifying directions, … When it reaches the synapse, the impulse can be blocked, changed from single impulse to repetitive impulses, integrated with impulses from other neurons to create complex types of impulses in successive neurons or transmitted directly without integration along with impulses from other neurons to the postsynaptic terminal.

For the formation of synapses, it is assumed that there is a specific chemical mechanism for mutual neuronal recognition. It appears that both pre- and postsynaptic surfaces have concordant clusters of specific molecules on certain neurons, like a key and a lock. This has been partly confirmed by experiments on fish and amphibians where, when their optic nerve is severed, the nerve fibres grow back and the neurons reach their target on an individual and differentiated basis, just like the original ones.

Concerning glia, the numerical ratio of glial cells to neurons, known as the Glia Index, in grey matter, the area where the cell bodies of the central NS are located is between 1 and 50. The glia owes its name to Wirchoff, who named it in 1948. Etymologically, it means "nerve glue". In the early 1980s, new methods of identification and cell culture, especially of astrocytes that form the myelin that facilitates the transmission of nerve impulses, began to reveal new functions for these cells. The other two types of cells found in the glial system are the oligodendrocytes that overlie neurons and the unmyelinated neural structures of astrocytes, whether they are myelinated axons or not. And the microglia cells that act, among other functions, as the defence system of the NS. Astrocytes are now considered to have multiple active missions in the maintenance of brain physiology. For example, they play a decisive role in the metabolism of glutamate and gamma amino butyric acid (GABA), both of which are NTs, the former being excitatory and the latter inhibitory to the central NS. Noremberg and co-workers demonstrated that glutamine synthetase, the enzyme responsible for glutamine formation, is found exclusively in astrocytes. Once synthesised, glutamine goes to the neurons where it becomes, arguably, the main source of GABA and glutamate in cells that use these substances as NT. Another highly important function of astrocytes is the so-called potassium spatial buffering, hypothesised in the mid-1970s by Kuffler et al. Some studies show that astrocytes have receptors for most

NTs. Although their function remains unclear, their presence suggests that astrocytes respond to changing conditions in the brain, with a versatility that may be similar or even superior to that of neurons.

The most important component of the NS is the brain, which is a somewhat complex system, and in the case of humans is made up of more than 10^{11} neurons in the cortex, forming a network of more than 5×10^{14} neural connections. A neuron can have up to 10^5 connections, although the average is between 5×10^3 and 10^4 connections.

In terms of how it works, the brain may be perceived as a system that performs tasks very differently to how they are performed by today's computers. Although the latter are very fast and efficient in certain information handling processes, there are highly complex tasks such as shape or pattern recognition and classification, which require too much time on today's more powerful computers. However, the animal brain turned out to be much better able to perform them successfully, sometimes even without apparent effort. For example, recognising a familiar face in a crowd of faces.

Although there are different types of biological neurons, the different parts of the neuron may be recognised in them, namely (See Fig. 2.2):

(a) A central body, called the soma, which contains the cell nucleus and is the integrating area for the information that reaches the neuron.
(b) An extension of the soma called the axon, the terminal end of which branches off and forms the area that transmits information.
(c) The dendritic tree, consisting of dendrites, which also originates in the soma, is highly branched and is primarily the information reception area.

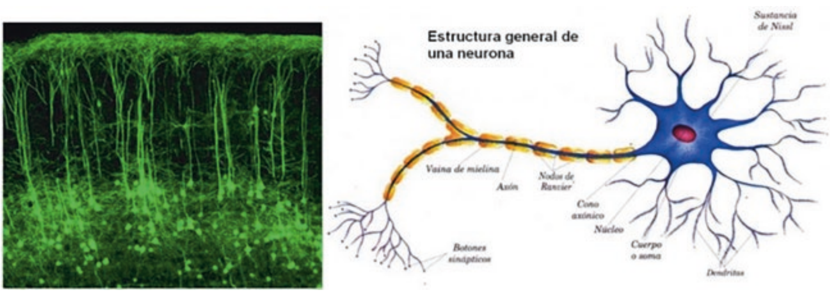

Fig. 2.2 Left, cerebral cortex section. Right, different parts of neuron typical

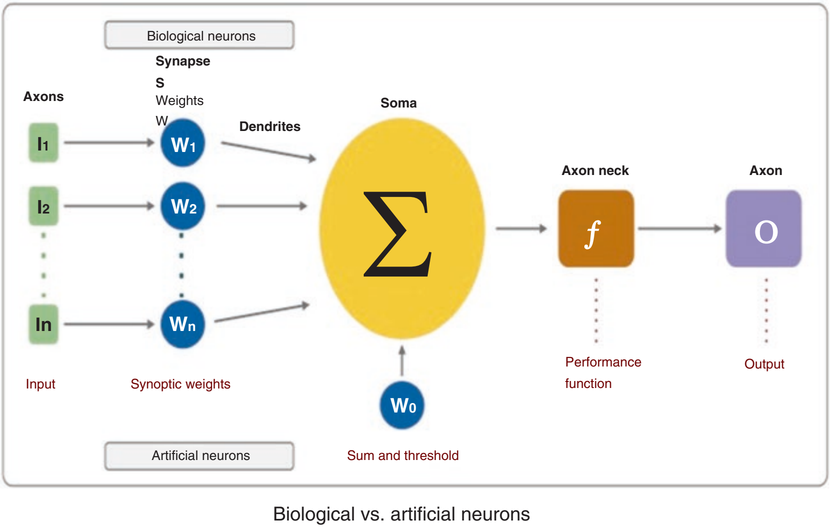

Fig. 2.3 Biological versus artificial neurons

(d) Synapses or areas of connection between neurons.
(e) The primary function of neurons is to transmit nerve impulses. These impulses travel throughout the neuron, beginning at the dendrites until they arrive at the axon endings, where they pass to another neuron via the synaptic connection. How living things respond to stimuli from the outside world, and how they learn about it, is directly related to the brain's neural connections. The ANNs provide a computational model to emulate this activity (See Fig. 2.3).

2.3.3 Artificial Neural Networks (ANN)

An ANN is a computational paradigm initially inspired by how the brain is assumed to function. The key element of this paradigm is its structure. ANNs consist of a number of processing elements or neurons working together to solve a specific problem. Current ANNs are based on the aforementioned mathematical neuron model, proposed in 1943 by McCulloch and Pitts. It modelled a simplified structure and functioning of brain neurons as devices with m inputs, a single output and only two possible states. Its components are:

- Neurons. They constitute the basic units of the network. Each neuron receives multiple different information from multiple inputs, processes the received information and produces a single output as a response that is transmitted identically to multiple downstream neurons. Similar to biological neurons, therefore, on one hand there is convergence, that is many inputs are converted into a single output. On the other hand there is divergence, that is the output information is transmitted to multiple neurons.
- Synapse or areas of connection between neurons and, similar to biological networks, information is transmitted unidirectionally, except in certain types of networks, for example Hopfield networks, where there is bidirectionality. Synapses may be of the following two types: Inhibitory, those that emit information that inhibits the neurons that receive it; that is, they decrease the level of action of the receiving neuron according to the information received. And excitatory neurons where exactly the reverse occurs.
- Synaptic weight. This key concept refers to the biological notion of the binding force of a synapse. A synapse is strong, or has a high connection weight, if the information it receives will contribute greatly to the new state produced in the receiving neuron and the response it produces. In ANN, synaptic weights are numerical values by which the signals received by the synapse are weighted. Thus, inhibitory synapses will have negative synaptic weights.
- Total input. Given that a neuron receives excitatory and inhibitory signals from multiple neurons, it becomes necessary to calculate the net effect that this set of signals or inputs will have. It is usually represented by NETi(t).
- Activation, which establishes the degree or level of excitation of a neuron. The activation level of neuron i at time t is denoted as Ai(t). Biological neurons are capable of storing a certain internal state of activation. The state of activation changes depending on the inputs received by the neuron, and the neuron's output or response depends on the state of activation. In ANN, an attempt has been made to model this operation by means of activation values.
- Neural response or output. It is the signal emitted by a neuron, represented by Oi(t), and indicates the signal emitted by neuron i at time t.
- Learning: It is the process by which neurons adjust their responses so that the behaviour of the network is as awaited. Basically, the synaptic weights are modified during the learning process. Later we shall dis-

cuss the rules governing this process of modifying synaptic weights in order to adapt the response emitted by the network, given specific input stimuli, to the correct response.

For years, neuroscientists' theories about how the brain learns were basically guided by Hebb's rule, which was often synthesised as follows: Neurons that fire together, wire together. That is, the greater correlation between the activity of adjacent neurons, the stronger the synaptic connections between them. With some modifications, this principle was able to account for certain limited types of learning and visual classification tasks. This rule however was much less workable with large ANNs that had learned from their mistakes. There was no direct target-oriented way for the neurons in the deep layers of the network to learn about and update themselves and make fewer mistakes from the errors discovered (Hertz et al., 1991). The Hebbian rule, in short, was a very narrow, particular and not very sensitive way of using the information obtained from the errors (See Fig. 2.4).

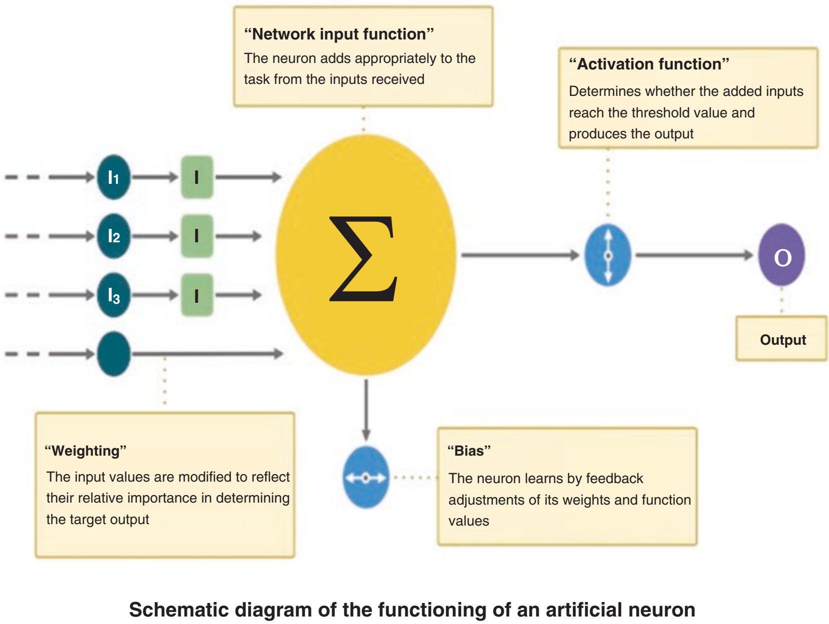

Fig. 2.4 Schematic diagram of the functioning of an artificial neuron

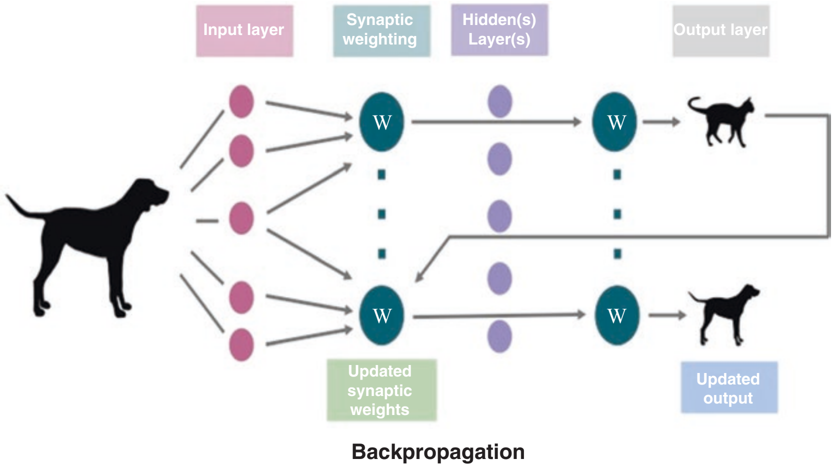

Fig. 2.5 Backpropagation diagram

The diagram of the operation of an artificial neuron is as follows: After receiving multiple inputs, each weighted by a factor or weight, that is a number expressing the importance assigned to that input, the neuron adds up the weighted inputs and produces, according to an activation function, an output signal, as shown in Fig. 2.5. From the 1960s onwards, it was clear that these neurons could be organised into a network with an input and an output layer, and that the network could be trained to solve certain kinds of simple problems. During training, the ANN established the best weights for its neurons to eliminate or at least minimise errors. However, it was clear until the 1960s that solving more complicated problems would require one or more layers of hidden neurons, that is neurons sandwiched between the input and the output. But it was not known how to effectively train ANNs with these hidden layers until 1986, when Hinton and colleagues revealed the backpropagation algorithm. This algorithm operates in two stages. One, forward, when it receives an input it infers an output that may be erroneous. Two, backwards; it updates the synaptic weights so that the output better matches a proposed target value.

In order to understand the process of backpropagation, consider a loss function that describes the difference between inferred and desired outputs in the form of a landscape with hills and valleys. When a network makes an inference from a given set of synaptic weights, it ends up somewhere in the loss landscape. To learn it needs to go down the slope, or gradient, into some valley, where the loss will be minimised as much as possible. That is, backpropagation is a method that updates synaptic weights for gradient descent. Essentially, the backwards stage of the algorithm calculates the contribution of each neuron's synaptic weight to the error and updates it to improve network performance. This calculation proceeds sequentially backwards from the output layer to the input layer, hence the name backpropagation. Doing this over and over again for desired sets of inputs and outputs will eventually lead to an acceptable set of weights for the entire ANN. However, it is widely believed that backpropagation is biologically implausible for several reasons. The first is that computers can execute the two-stage algorithm without difficulty; however, this would not be easy for networks of biological neurons. The second is the weight transport problem as computational neuroscientists call it, that is the backpropagating algorithm copies or transports information on the synaptic weights involved in an inference and updates those weights for greater accuracy. But in a biological network, neurons only see the outputs of other neurons, not the synaptic weights or the internal processes that shape that output. Now, from the point of view of a neuron, it is good is to know its own synaptic weights, not the set of synaptic weights of another neuron. For a learning rule to be biologically plausible, it has to respect this constraint: Neurons only access information from neighbouring neurons. However, backpropagation may require input from remote neurons. Consequently, as Bengio pointed out, if backpropagation is taken at face value, it does not seem possible for brains to compute (See Fig. 2.5).

As a matter of fact, the major difference between an ANN and a conventional computer is that ANNs, to some extent, process the input information to produce an output or response. It is not the blind and automatic application of an algorithm.

2.3.4 ANN Typology

There are different types of ANNs that use very different design philosophies, learning rules and construction of response functions, as shown in Fig. 2.4.

An initial cataloguing is conducted according to the route followed by the information within the network. In this regard, we may distinguish between the following two types of ANNs: One is forward-fed. Where information flows in only one direction: Backwards and forwards.

Classically, these networks are organised in layers. Each layer contains a set of neurons that receive synapses from neurons in the previous layer and send outputs to neurons in the next layer, without synapses between neurons in the same layer. In this type of network, there is at least one input layer, formed by the neurons that receive the input signals to the network, and an output layer, of one or more neurons that emit the network's response to the exterior. Between these input and output layers there are one or more intermediate layers. In networks constructed in this way, it is evident that information can only move in one direction: From the input layer to the output layer, passing through each and every intermediate layer only once. When information reaches a certain layer, that is when the neurons of a certain layer receive information from the previous layer via their input synapses, they all calculate their new activation state in parallel and send their response to the neurons of the following layer, which will repeat the same process. Information processing in these networks is performed within the same predetermined time for all possible input signal configurations. The fact that there is no connection between neurons in the same layer means that there are no waiting times in which neurons are interacting with each other until the entire layer is stabilised. These networks are therefore able to make rapid calculations. The other type of ANN is feedback ANN. In this type of network, neurons are not grouped into layers. Each neuron is connected to every other neuron and thus, when information is input into the network, each neuron will have to calculate and recalculate its state several times, until all neurons in the network reach a stable state. This stable state is one in which no changes occur in the output of any neuron. Without changes in outputs, the inputs of all neurons will also be constant, which means they will not have to change their activation state and response, thus maintaining a stable global state. A neuron in the network that directly receives a signal from the outside may initially take on a certain state of activation and accordingly emit a response that will act on other neurons that in turn may influence the initial neuron with their outputs, forcing it to take on a different activation state and so on until they all become stable. In these networks, there is no way of knowing how long it will take to reach a stable state. Moreover, some input stimuli will cause a stable state to be reached in less time than others. Hopfield networks, which are very different from other types of ANNs, are the most prominent example of this type of monolayer feedback networks.

There are multilayer networks that use feedback at one or more layers. In these networks, there are synapses between neurons in the same layer. Networks with this architecture are called multilayer networks with lateral feedback. In these networks, a neuron can feed back on itself. Indeed, these feedback networks most closely emulate the structure of the human brain, where feedback phenomena are abundant and essential.

2.3.5 Relationships Between Inputs, New Status and Response in ANN: Activation and Transfer Functions

As we have previously mentioned, all neurons receive information at their input synapses, change their activation state according to these inputs and emit a response. However, it has not yet been explained how these processes are conducted. In this sense, the first operation performed by a neuron is to compute the incoming signals and calculate the total input NETi(t) that it receives at time t. This process is performed according to the signal propagation rule, which is expressed as follows:

$$\mathrm{NETi}(\mathrm{t}) = \sum Wij\ Oj\ (\mathrm{t}-1).$$

NETi being the weighted sum of all signals that arrive at neuron i, Oj, each of them weighted by the associated connection weight, Wij. When a given signal Oj arrives at neuron i from an inhibitory synapse, the weight Wij associated with that synapse will be negative; therefore the result of the product Wij×Oj will add a negative element to the sum.

On the other hand, the activation and transfer functions are responsible for defining the new activation state Ai and the neuron's response Oi, based on the set of inputs Oj that the neuron receives at a given time and, sometimes, on its previous activation state. More specifically, the activation function is a mathematical function that relates the input information to the next activation state of the neuron. The network designer must decide how many possible activation values a neuron can take.

In so-called bounded models, the activation value of a neuron i may be any value within a range of continuous values. In other, so-called unconstrained models, there is no limit to activation values. In simple models, the activation value is discrete; the neuron can only have two states: Activated or deactivated, these states being identified by the pair of values [0,1] or [-1,1], respectively.

When designing a network, in addition to establishing what the activation values of each neuron will be, the activation function with which each neuron will process its inputs to elaborate its new activation state is also decided. Information reaches a neuron via its input synapses, both excitatory and inhibitory. Each synapse, as we have seen, has a certain connection weight associated with it, which will be negative in the case of inhibitory synapses. The activation function will then act on the input signals, the synaptic weights associated with each input, and sometimes on the activation value that the neuron had at the time the signals were received. Formally, as shown in the following formula:

$$\mathrm{Ai}(\mathrm{t}) = FA\left[\mathrm{Ai}(\mathrm{t}-1), Oj(\mathrm{t}-1), Wij\right] \text{or which is the same :}$$

$$\mathrm{Ai}(\mathrm{t}) = FA\left[\mathrm{Ai}(\mathrm{t}-1), \mathrm{NETi}(\mathrm{t}-1)\right].$$

The activation value of neuron i at time t will depend, according to how the activation function FA is set, on:

- The activation state of neuron i at time -1.
- Of all the inputs On received from the j neurons of the previous layer at time t-1.
- Of the connection weights Wij associated with the different synapses through which each signal Oj arrives at neuron i.
  A very simple activation function that does not consider the previous activation state is the following: Ai(t)=Sj Wij(t-1). According to this activation function, the current state Ai(t) is the sum of all received signals Oj weighted by the synoptic weight Wij associated with the synapse through which it arrives. If the set of input stimuli arriving at a given layer is represented as a vector, and a matrix representation is maintained for the synaptic weights, the calculation of the new activation state of all those in the layer is obtained by a simple matrix product. Indeed, supposing that the values of the stimuli received by a network are: e1=1, e2=0, e3=2, e4=1, e5=2, and that the weights matrix is

$$\begin{vmatrix} 2 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{vmatrix}$$

> then multiplying the vector of the values of the received stimuli: [1 0 2 1 2] by the above matrix gives the following activation states: [2 5 4], from which the neuron will make its response. The simplicity of the activation function used in the example is evident as it does not even consider the previous state of the neuron. Sometimes it is interesting to include a factor representing a certain fraction of the previous activation value in the activation function, for example Ai(t)=1/2Ai(t-1) + j Wij Oj(t-1).
> With this activation function, even when the inputs are zero, the neuron will maintain a certain state of activation that will decrease over time. In each time period, the network will change to an activation value that is half of the previous activation value. These activation functions ensure that individual neurons do not vary too abruptly in their state and thus in their response. What actually happens is that the neuron thus has a certain capacity for self-feedback, both in the excitatory and the inhibitory sense.

A widely used type of activation function is the threshold function. As we have already seen, there are some networks where each neuron can only have two states: Activated and deactivated; these networks use a threshold activation function, which is defined as follows: Ai(t)=1, if j Wij×Oj(t-1) > threshold; otherwise =0.

As can be seen, the neuron will be activated, that is set to value 1, if and only if the weighted sum of its inputs exceeds a certain threshold, which is decided by the network designer. An interesting type of activation function, called stochastic function, includes a small factor of random value in the function, which represents disturbances or possible noise that may arise in actual neural networks.

In short, there are many possibilities for designing activation functions and it is up to the designer to determine the activation function or functions to be used by the network. On the other hand, the transfer function, TF, provides the output value or response Oi that neuron i will emit depending on its activation value Ai at that moment. That is, Oi(t)=TF[Ai(t)]. In the same way that multiple activation functions have been defined, there are also multiple transfer functions. The simplest is the Identity function, whose output is the activation value of the neuron, that is Oi(t)=Ai(t). This function is typically used in a threshold activation function, so that the Oi output will be 1 or 0 depending on whether the neuron is activated or deactivated, or 1 and -1.

There are also threshold transfer functions, where the output response will either be 1 or 0, or 1 and -1, depending on whether the activation level, which in this case will be a continuous value, exceeds a certain threshold.

Oi(t)= 1 if Ai(t) > threshold; and 0 otherwise.

We may also consider an activation function that takes into account the activation value of the neuron at the previous moment to calculate the new state; the neuron thus receives feedback. The activation state will take values within a continuum, but the output Oi calculated by a threshold transfer function will be discrete, taking only one of two possible values 0,1 or -1,1. The transfer function where the response Oi(t) grows according to the growth of the activation state Ai(t), until it reaches a point where the response stabilises, such that further increases in activation do not generate larger responses, is called the saturation function. That is, increasing the activation level beyond that which produces the maximum response has no effect on the response. The most interesting transfer function is the sigmoid function. This function produces continuous outputs, within the range [0,1], proportional to the level of activation of the neurons, but maintains the saturation levels, a maximum at which it outputs 1, and a minimum at which it outputs zero. That is, when the activation level exceeds the maximum saturation threshold, the output will remain 1, and similarly, activation levels below the minimum saturation threshold will produce 0 outputs. This sigmoid function, in the form of an S-shaped Squashing Function, is: Oi(t)=1/[1/1+exp{-Ai(t)}].

Another transfer function that maintains an upper and a lower saturation limit is the hyperbolic tangent function. In this function, the output Oi(t) varies proportionally to the activation level, but within the range -1,1. The expression of this function is: Oi= exp{Ai(t)}-exp{-Ai(t)}/ exp{Ai(t)}+exp{-Ai(t)}.

In spite of the large number of activation and transfer functions that can be used, there is a classification criterion that divides them into two main groups: Linear and non-linear functions. The linearity or non-linearity of these activation and/or transfer functions is the most defining characteristic of a neuron's behaviour.

Linear neurons, as the name suggests, are neurons whose output Oi is linearly dependent on their inputs On, that is it is proportional to them. These neurons have linear activation and transfer functions; therefore the composition of the two functions leads to another linear function that will govern the production of responses according to inputs. In linear neurons, the response is the result of applying a linear function to the sum of the inputs Oj, each weighted by the synaptic weight Wij associated with the synapse through which each Oj arrives. Therefore, the responses of linear neurons are not bounded and can take any value.

The first ANNs were of this type and their linearity presented the following two major problems: One, a lack of persistence. Clearly, with a linear function, changes in the inputs inevitably produce changes in the outputs regardless of whether the output Oi has a high or low value. Thus, very small changes in the inputs can produce significantly large fluctuations in the response. This lack of stability in a neural response is an undesirable characteristic, as it produces constant instability in the response of the network as a whole. Two, lack of simultaneous suitability for large and small signals. If the function resulting from the composition of the activation and transfer functions has amplifying effects on the input signals, then small input signals will not be lost but will elicit a response from the neuron. But in this case, moderate entry signals will produce very high responses. On the other hand, if the neuron is poor at amplifying, that is if it produces moderate outputs to medium input signals, then weak input signals will be lost. It is impossible, as we shall see, to make a linear neuron adapt its signal to both large and small signals.

Clearly, if you have a network with a linear activation function and the transfer function is the identity, you have: Oi=AI=jWijOj.

The set of responses Oi of the neurons in a layer form a vector, which is simply the product of the vector of signals Oj arriving from the previous layer, through the synaptic weight matrix Wij linking the two layers. If more intermediate layers are added, the process is repeated and the final output of the network will once again be a vector of On signals, resulting from successive matrix products. In other words, the initial vector of inputs is multiplied by the matrix of weights Wj associated with the first layer, and the result of this product is, in turn, a vector that is multiplied by the matrix of weights Wij associated with the second layer, and so on until the output vector of the network is obtained. In this process, all the intermediate products of the Wij matrices associated with the different layers may be replaced by a single matrix product which would be the result of the multiplication of all the intermediate matrices. Thus, a network with only two layers, in which the weight matrix associated with the second layer is the resulting matrix, would produce exactly the same results as the multilayer network above.

However, in non-linear neurons either the activation function or the transfer function or both are non-linear functions so that the composition of these functions will also be non-linear, resulting in the responses of the neurons not being a linear function of the inputs. It is easy to observe that, for example, the threshold, sigmoid and hyperbolic tangent functions are non-linear.

Non-linear neurons produce bounded responses. If it is a simple threshold neuron, one whose activation or transfer function is of the threshold type, its output is discrete, thus eliminating the problems of response fluctuation and mismatch to large or small signals. Neurons with a sigmoid or tangent hyperbolic function or, in general, a transfer function with an upper and lower saturation limit, allow responses to vary sensitively to changes in inputs only when the neuron is halfway between the two saturation levels. But when the neuron reaches one of the saturation levels, small fluctuations in the input signals will not alter the response, which will remain stable at that saturation level. This effect is highly desirable, as it ensures a certain persistence of the neural response. It is as if the neuron is blocked to some extent, during the emission of the response, to the effects of new input signals, which are not very different from the current ones.

2.3.6 ANN Construction Methodology

Obtaining a solution to solve an ANN-based problem involves the following steps:

1. Define the input. The input problem, as far as an ANN is concerned, consists of a series of numbers. Such inputs may be, in visual pattern recognition systems, a two-dimensional array of numbers representing the pixels of an image. In auditory recognition systems, for example speech, a two-dimensional array of numbers representing sound, in which the first dimension, sound parameters, for example, are frequency components, and the second dimension represents different moments in time. And, in random pattern recognition systems, an n-dimensional array of numbers represents the pattern input.
2. Define the network topology. For example, the layers of neurons and the connections between neurons. When assembling the neural network, the architecture of each neuron consists of:

(a) Multiple inputs where each input is connected either to the output of another neuron or to one of the input numbers.
(b) Usually a single output that is connected either to the input of another neuron, which is usually in a higher layer, or to the final output. To assemble the first layer of neurons, the following is required: On one hand, to create N0 neurons in this layer. For every one of these neurons, each of the multiple inputs of the neuron must be connected to points, for example numbers, in the problem input. These connections may be determined randomly or by means of an evolutionary algorithm. On the other hand, an initial synaptic strength must be assigned to each connection created. These values may start out equal, may be randomly assigned or may be determined, as we shall see below, in other ways.
(c) To assemble the additional layers of neurons, the following must be done for each layer: On the one hand, Ni neurons of the i-layer must be created. For every one of these neurons, each of the multiple inputs of the neuron must be connected to the outputs of the neurons in layer i-1 as will be seen below. Moreover, an initial synaptic strength must be assigned to each connection created. These values may start out the same, be randomly assigned or be determined in another way. Finally, the outputs of the m-layer neurons are the outputs of the ANN. Regarding the recognition tests of the workings of each neuron; that is, once the neuron is assembled, it does the following in each recognition test:

    i. Each neuron's weighted input is computed by multiplying either the output of another neuron or the initial input, that is to say, the input to which the neuron is connected by the synaptic strength of that connection.
    ii. All these weighted inputs from the neuron are added together. If this sum is greater than the activation threshold of this neuron, then this neuron is considered to be activated and its output is 1. If not, its output is 0.
    iii. For each recognition test, for each layer, from layer 0 to m, for each neuron in the layer, its weighted inputs must be added, each weighted input = to the output of another neuron, or it is an initial input; that is, the input to which that neuron is connected multiplied by the synaptic strength of that connection. If this sum of input values is greater than

the activation threshold of this neuron, set the output of this neuron = 1, otherwise set equal to 0.

3. Train the ANN with examples of the problem. For this purpose, repeated recognition tests have to be carried out on sample problems. After each test, the synaptic strengths of all interneurons must be adjusted to improve the performance of the tested ANN. This training should be continued until the accuracy rate of the ANN stops improving, for example, when it reaches an asymptote.
4. Run the trained ANN to solve new examples of the problem. For that, the designer of such an ANN needs to determine the following at the outset: What the input numbers represent; the number of layers of neurons; the number of neurons in each layer, as each layer can have different numbers; the number of inputs for each neuron in each layer; the number of interneurons, which can also vary from neuron to neuron and from layer to layer; and the actual wiring, for example connections. For each neuron in each layer, this consists of a list of other neurons, whose outputs constitute the inputs of this neuron. This is an essential area of design. To do this, there are a number of possible ways ranging from randomly hard-wiring the ANN, to using the network designer's judgement, to using an evolutionary algorithm.

The initial synaptic strength must also be established, for example the weights of each connection. There are different ways of doing this, including the following: Give synaptic forces the same value or give them different random values; use an evolutionary algorithm to determine an optimal set of initial values or, again, use the system designer's judgement to determine the initial values. Similarly, the activation threshold of each neuron must be established.

On the other hand, the output may consist of: The outputs of the m-layer of neurons; the output pertaining to the output of an individual neuron, whose inputs are the outputs of m-layer neurons; a function of, for example, a sum of the outputs of m-layer neurons; and another function of the outputs of multilayer neurons.

How the synaptic strengths of all connections are adjusted during the training of this ANN is a fundamental design decision and is subject to a great deal of research and discussion. In any case, there are several ways of doing this: On one hand, it may be done by increasing or decreasing each synaptic strength, in each recognition trial, by a fixed, usually small, amount so that the ANN output is closer to the correct response. One way to do this is to try to both increase and decrease the strength, and see which performs better. This may be time-consuming, so there are other methods for taking concrete decisions on whether to increase or decrease each synaptic strength. In this sense, there are statistical methods to modify synaptic strengths after each recognition test, so that the ANN's performance on that test is closer to the correct answer. Note that ANN training will work even if the answers to the training tests are not correct. This allows the use of real-world training data that may contain an inherent error rate. One of the keys to the success of an ANN-based recognition system is the amount of data used for training. Normally, a large quantity is needed to obtain satisfactory results. As with human learners, the amount of time an ANN spends learning its lessons is a key factor in its performance.

In addition to the above, many variations are possible, including the following: On one hand, there are different ways to determine the topology of the network, especially the interneuron wiring may be randomly assembled using an evolutionary algorithm. On the other hand, there are different ways of establishing initial synaptic strengths. Likewise, the inputs of i-layer neurons need not come from the outputs of neurons in layer i-1. Conversely, the inputs to the neurons in each layer may come from any layer below or from any layer. Equally, there are different ways of determining the final outcome.

The method described above results in an all-or-nothing activation, that is to say, 1 or 0, called non-linearity. There are other non-linear functions that may be used. Normally, a function that goes from 0 to 1 in a fast but more gradual manner is always used. Also, the outputs may be numbers other than 0 and 1.

It is not in vain to insist that choosing a good method of adjusting synaptic forces during training is a fundamental design decision. The above scheme describes a synchronous ANN, where each recognition trial takes place by computing the outputs of each layer, starting from layer 0 to layer m. However, in a truly parallel system where each neuron operates independently of the others, the neurons can operate asynchronously, that is, independently of each other. In an asynchronous approach, each neuron is constantly scanning its inputs and is activated each time the sum of its input values exceeds the threshold or whenever specified by the output function.

2.3.7 *Learning in ANN*

Just as it is impossible, as demonstrated in Chap. 1, to provide a formal definition of intelligence, there is a similar formal definition of learning. It was authored by Wiener (Wiener, 1964) and even informally stated as follows: An organised system may be defined as one that transforms a certain input message into an output message, according to some transformation principle. If such a principle is subject to certain performance validity criteria, and if the transformation method is adjusted so that it tends to improve the performance of such a system in accordance with these criteria, then the system is said to learn. This definition is valid for both ontogenetic learning, which is learning that takes place in an individual, human or not, and phylogenetic learning, or learning that affects the species.

Learning in neural networks is defined as the process by which the network modifies its responses to its inputs in order to gradually adapt to the operation that is considered correct. There are two types of learning in ANN. The first, called supervised learning, consists of repeatedly presenting the network with patterns of input stimuli from a training set. This game consists of a stimulus-correct response pattern pairs and must be chosen carefully. Each pair has a de facto name. All the information that the network needs to learn must be represented in the training set. The network's response to each pattern is compared with the correct response to that pattern, and by virtue of this comparison the synaptic weights are automatically readjusted using learning procedures discussed later. The synaptic weights are readjusted to ensure that, given the input pattern, the network gives the correct response. When the network responds well to an input pattern, it moves on to the next pattern in the training set and repeats the same procedure. After completing the last pattern of the training set, we start once again with the first one, as the weights have continued to change.

In the second approach, called unsupervised learning, the network is not told what is the correct response, that is there is no comparison between the network's response and the desired response. Moreover, there is no external influence on the network in this learning model, as it is not informed of whether a result was correct or not; the network is only provided with large amounts of data with which it can build its own associations. A much larger number of input patterns are therefore needed in training so the network may correctly adjust its synaptic weights. What is done in this learning model is actually asking the network to capture some of the characteristics of the input data by itself. In this type of learning, the neurons are expected to organise themselves by learning to grasp the regularities of the input data without any external criteria or aids to direct this self-organisation. Certainly, this type of learning is the most interesting one and it is currently used in the most advanced ANNs.

After the training phase comes the execution phase, during which the network will be asked to respond to stimuli different from those presented during the training phase. Based on the examples learned from the training set, the network should be able to generalise and give correct responses to new stimulus patterns. In other words, once learning is complete, a network is able to generalise; that is, given inputs similar to those in the test set, it will produce correct outputs. It should be noted that it is very difficult to test the generalisability of a network without using large amounts of data.

The most important thing about a neural network is undoubtedly the learning procedure or rule. This is precisely what makes a neural network different from a conventional program and gives it its special characteristics and abilities. As the network learns, its synaptic weights are modified. This modification may lead a weight to become zero, to become non-zero, or to change from a positive to a negative value or vice versa. When a weight becomes zero, the effect is the same as if the connection ceases to exist. Learning rules or procedures define how to change synaptic weights.

Over time, different learning rules and methods have been proposed, among which the best known are the following: The Hebbian Learning Rule of 1949; and the Widrow-Hoff Learning Rule or Delta Rule of 1959, generalised in 1974, also called the gradient backpropagation rule. This rule was formalised and used operationally by Rumelhart and colleagues, Le Cun and Parker independently, in 1985. From then on, a multitude of learning procedures was proposed that improved on earlier rules. Specifically, within unsupervised learning, we have Kohonen's networks of 1982. Also called Kohonen's feature map or Kohonen's self-organising map, SOM, the main characteristic of these networks is that they perform a topological representation, that is a classification, of the set of learning patterns. The resulting learning map shows the relationship between the input patterns. The closeness or remoteness between the representations of two patterns indicates the degree of similarity between them. The network is composed of two layers called the input layer and the competition layer, and as mentioned above, its form of training or learning is unsupervised. That is to say, Kohonen's map finds the underlying organisation of the learning set. Patterns are classified according to the unit that is activated in the competition layer. If two patterns are similar they will be classified in the same unit in any of their neighbourhood. Once the learning is over, it is possible to classify the patterns, to see how they are grouped and to give a graphical idea of the learning set. The advantage of Kohonen's maps is that they can make a one or two-dimensional representation of a set of patterns of any number of components, n-dimensional, making it possible to visually capture their structure and characteristics.

A simple but real-life example will enable a better understanding of the discourse regarding ANNs (Ríos, et al., 1991). The goal is to design a simple neural network that warns the pilot of a twin-engine aircraft about irregular situations. This entails the network monitoring certain instruments and indicating to the pilot whether the flight conditions are suitable. It will not report anything specific but simply warn of danger when it detects an irregularity, without specifying which one. This means the pilot won't need to constantly monitor the instruments as the network will do it for him.

Suppose you have one of the many ANN building tools available in the market. The tool is then given the following characteristics: Number of network layers; number of neurons in each layer; activation and transfer functions of neurons; and complete training set. It would be sufficient to use a simulation tool for forward-fed networks that are trained using the generalised Delta rule.

The parameters that will be taken into account to see if everything is working properly are:

A. Flight situation. A distinction is made between the following four situations: Take-off, automatic flight, that is constant speed without changes in altitude and trajectory, flight with manoeuvres, and landing.
B. Landing gear position.
C. Fuel level.
D. Temperature of each engine.
E. Revolutions per minute (rpm) of each engine.
F. Altitude.

In this case, it is fairly straightforward to represent the information from the instrument readings. It is assumed that the network receives all readings in the form of renormalised numbers to fit the range of input values of the neurons in the network, as will be shown below. It is therefore necessary to know the range of permissible values for each of the

Table 2.1 Valid values for twin-engine parameters

Parameter	*Range*	*Take-off*	*Automatic flight*	*Manoeuvring flight*	*Landing*
Landing gear	Outside-inside	Outside	Inside	Inside	Outside
Fuel level (litres)	0–400	>60	>40	>50	>15
Engine temperature	From -5 °C to 95 °C	10–60 °C	20–70 °C	20–80 °C	20–90 °C
Engine rpm	0–5000	500–3000	500–3000	700–4000	1000–4000
Altitude (m)	0–3000	0–500	500–2000	500–2000	0–500

parameters in each of the flight situations. It is also known that the values considered valid for parameters B, C, D, E and F are those given in Table 2.1.

It is then necessary to represent the output values of the instruments, which shall be within the ranges given in Table 2.1, in a way that the network can understand them. That is, these values must be renormalised within a range 0–1 in which the input values of the network are represented. Thus, for:

A. Flight situation. Two neurons will be sufficient to represent the four possible flight situations. Inputs to these neurons will take only two values: 0 or 1, thus making: 00 take-off; 01 automatic flight; 10 flight with manoeuvres; 11 landing.
B. Landing gear. A single 0|1 dichotomous input neuron is sufficient for this: 0 undercarriage outside; 1 undercarriage inside.
C. Fuel level. Only one neuron is required to receive continuous value inputs within the range 0–1. For this purpose, the possible fuel level values (0–400 litres) must be renormalised within the input range. Logically, it is enough to divide the fuel level by 400, and the result will be the grid input.
D. Engine temperature. Since the aircraft has two engines, two input neurons are needed, one linked to each engine. The input from these neurons shall be continuous in value. The temperature of each engine, which ranges from -5 to 95 °C, must be renormalised in order to be expressed on a scale of 0 to 1. To do this, 5 is added to the temperature given by the instrument at any time and the result is divided by 100. The value thus obtained is the one provided to the corresponding neuron.

E. Revolutions per minute of the engines. Same procedure as in the previous case. There will be an input layer neuron associated with the rpm of each engine. The input to each of these neurons will take a continuous value obtained from the renormalisation of the associated number of engine revolutions, within a 0–1 scale.
F. Altitude. An input layer neuron associated with this parameter is needed here. The input to this neuron will take continuous values obtained from the renormalisation of the altitude supplied by the aircraft altimeter, within a range of 0–1.

After representing the input information, it is obvious that the network output will be a single signal indicating whether everything is going well or not. A single neuron in the output layer, whose response takes values of zero or one, will therefore suffice. Presumably you have set me to zero which causes the Paris indicator light to switch on, whereas an answer of one, which is associated with a situation of normality, does not cause the Paris indicator light to switch on.

Until now the easy part, since it follows logically from the type of problem and how the input and output information are represented in the network. Then comes the tricky part, the network design. We know that nine neurons are needed in the input layer: 2 for flight status; 1 for landing gear status; 1 for fuel level; 2 for temperature of each engine; 2 for revolutions per minute of each engine; and 1 more for altitude. And 1 neuron is needed in the output layer.

In the middle layer, five neurons may be placed initially, about half the total number of neurons in the input and output layers. It may be of interest later to test with a higher number of neurons in this layer and see if better results are achieved, such as faster training, fewer errors, and better generalisation of the patterns learned during training. The Fig. 2.6. represents what has just been said.

A continuous transfer function is also needed, as many of the input values are continuous. Furthermore, it is not desirable for the neurons in the network to have linear behaviour. A sigmoid function may therefore be used, since this function allows the neuron to vary sensitively only at medium levels of the input signal, but falls off very quickly at high (1) or low (0) saturation levels, which is interesting since the output neuron is intended to give values of 0 or 1 (See Fig. 2.6).

Once the network has been designed, it moves on to the training phase. For this, it is necessary to build a large set of correct input-response

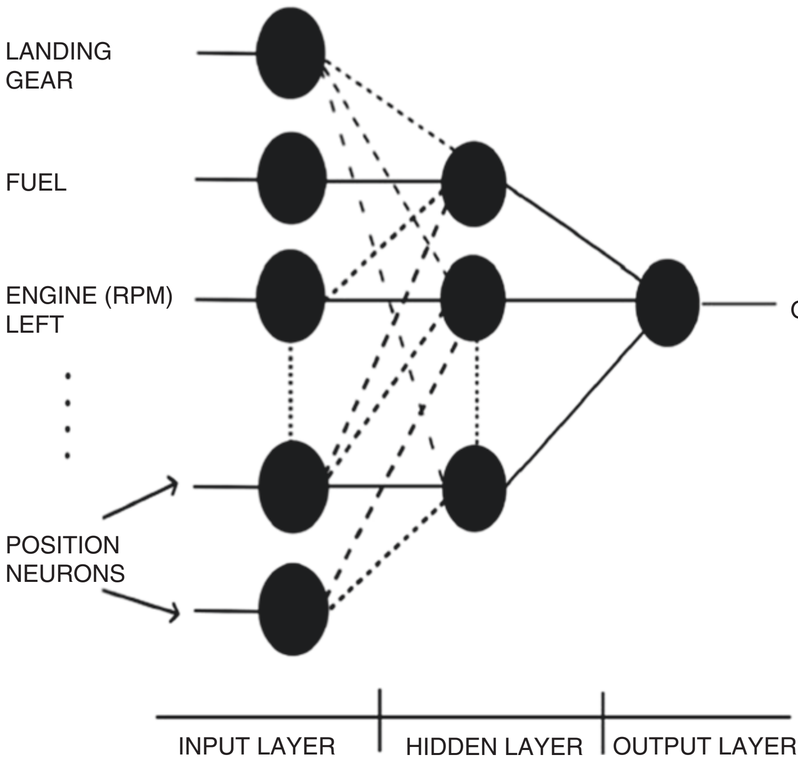

Fig. 2.6 ANN figure of twin-engine flight status assessment

patterns to be used as training facts for the network. Within this training set, it is essential to design events where the network response must be 0, (indicating problems). Additionally, a 0 response event should be included, where the problem is caused by only one of the parameters, for every parameter in each flight situation. In other words, input patterns should be included in the training set in which all -1 parameters have a correct value. That way, it would simultaneously learn that a value outside the allowed range in just one of the parameters is enough to cause problems. A real example of the training set could be seen in Table 2.2:

Given this input pattern, the output value should be 0, that is it indicates that something is wrong, even if only one parameter is out of range, in this case, the right engine temperature. Finally, all that remains is to test the

Table 2.2 Neuron entry information and value

Entry information	*Neuron*	*Entry value*
Flight situation	0	0 (automatic flight)
Flight situation	1	1
Landing gear	2	1 (retracted)
Fuel	3	70 litres
Left engine temperature	5	50 °C
Right engine temperature	5	90 °C
Left engine rpm	6	2500 rpm
Right engine rpm	7	2500 rpm
Altitude	8	700 m

network. This may be done by providing random instrument readings and observing what happens, or by taking actual flight data and seeing how the network behaves. However, the most useful thing to do is to reserve some training patterns, approximately 5–10%, and use them to test the network once it has been trained with the rest of the patterns.

If the network passes the test, then more rigorous tests may be conducted. The level of demand of network performance will depend on the end use of the network. Since it involves the safety of human lives, a flight status check network for a twin-engine aircraft should be tested very thoroughly and rigorously.

Last but not the least, as we shall see in Chap. 7, ANNs today have surpassed world chess and Go champions, as well as in other domains where humans were once unbeatable.

2.4 Identification of Strategic KBS Applications

KBS applications, even if successful and technically advanced, may fail to impress and win top management support if they do not yield a competitive advantage to the firm in the long run. A KBS application, if identified as of strategic benefit to the firm, will (with proper communication) be better placed to win sustained management support and resources even if it is technically complex.

2.4.1 Value Chain and Value Activities

Potentially strategic KBS applications can be determined by analysing the following types of activities and links within the value chain of the company:

High value activities: Activities which create high value to the company (in the context of its overall competitive strategy) are obvious first places to look for potential KBS projects. The knowledge processing occurring within high value activities should be analysed carefully to determine whether it may be possible to leverage that knowledge with a KBS. For example, computer configuration is a high value activity for Digital, and Digital's first large KBS project (the XCON system) was designed for that activity.

Activities interfacing with high value activities: Activities which either influence or are influenced by high value activities are sources for potential KBS projects. Such activities are important, but frequently overlooked sources of value to a company. For example, within Digital the activities of selling and manufacturing computers respectively influence, and are influenced by the high value engineering activity of configuration.

Junctions: Activities which serve as a junction for many linkages across different value activities can also lead to strategic KBS projects. Though such activities may not obviously appear as high value activities, they are often areas for intense knowledge processing. Bottleneck activities: Bottleneck activities are undesirable because they slow down the generation of value in the value chain. KBSs can potentially be used in such activities to alleviate the bottlenecks. Though many high value activities are bottlenecks, the two need not always be the same. For example, invoice 17 processing can be a bottleneck (but not a high value) activity within a particular firm.

Boundary linkages: Linkages between value activities across boundaries, both external (such as those between a firm and its suppliers or customers) and internal (such as those between different sub-divisions of the same firm), are usually associated with intense knowledge flows. KBSs can potentially be used to facilitate the flow of knowledge across these boundary linkages. For example, a KBS can be used by a supplier to get "intelligent" assistance while ordering products from the company. Such a KBS can facilitate the flow of knowledge about a company's products across to its customers.

2.4.2 Core Competencies and Capabilities

The core competencies of a corporation are loci of intense knowledge processing as they represent the coordination and integration of a diverse set of skills. From a knowledge management perspective, a company is able to leverage its core competencies if it can:

(a) identify its set of core competencies,
(b) collect the necessary (diverse) knowledge to build its core competencies,
(c) integrate and coordinate the knowledge elements related to its core competencies, and
(d) disseminate knowledge about its core competencies to all relevant parts of the organisation.

KBSs can play an important role in steps (b), (c) and (d).

KBSs can facilitate the transfer of diverse knowledge elements from across the company to the "core competency center" within the company. This is possible with either KBSs targeted at the individual (or group) level (collect the expertise of several experts) or the organisational and knowledge links level (share expertise across the organisation). KBSs can help with the integration and coordination of knowledge elements related to the core competency. This is especially useful if numerous, complex knowledge elements need to be aggregated and managed. Specific integrated knowledge about specific core competencies can also be distributed and shared within the organisation with the help of KBSs (targeted at either the individual or the organisational levels). While core competencies are focused on the integration and coordination of knowledge at specific centres within the corporation, capabilities are concerned with the development of skills across different activities in core business processes. From a knowledge management standpoint, a company is able to build strong capabilities if it can

(a) identify its set of core processes and their constituent activities, and
(b) share relevant knowledge across activities within core processes (and across core processes if required). KBSs targeted at the organisational and knowledge links levels can help with the sharing of critical knowledge within (and across) core processes and help the organisation build a set of core capabilities.

References

Barreiro, C., Barreiro, J. M., Lara, J. A., Lizcano, D., Martínez, M. A., & Pazos, J. (2019). The Third Construct of the Universe: Information. *Foundations of Science*. https://doi.org/10.1007/s10699-019-09630-7

Borrajo, D. (1993). *Inteligencia Artificial: Métodos y Técnicas*. Editorial Ramón Areces.

Gómez, A., Juristo, N., Montes, C., & Pazos, J. (1997). *Ingeniería del Conocimiento*. Editorial Centro de Estudios Ramón Areces.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley Publishing Company.

Kanigel, R. (1991). *The Man Who Knew Infinity: A Life of the Genius Ramanujan*. Charles Scribner's Sons.

Kipling, R. (1992). If—(Brother Square-Toes-Rewards and Fairies). In *The Complete Verses* (pp. 472–473). Kyle Cathie Limited.

Longfellow, H. W. (1861). *The Open Window*. University of Michigan: Humanities Text Initiative.

Post, E. L. (1947). Recursive Unsolvability of a Problem of Thue. *The Journal of Symbolic Logic, 12*(1), 1–11.

Prerau, D. S. (1990). *Developing and Managing Expert Systems. Proven Techniques for Business and Industry*. Addison-Wesley Publishing Company, Inc.

Ramanujan, S. A. (2012). *Notebooks* (2 Vols.). Tata Institute of Fundamental Research.

Ramón y Cajal, S. (1960). The Structure and Connexions of Nervous System. In *Nobel Lectures: Physiology or Medicine, 1901–1921*. Elsevier Science Publishing Co.

Ríos, J., Pazos, A., Brisaboa, N. R., & Caridad, S. (1991). *Estructura, Dinámica y Aplicaciones de las Redes de Neuronas Artificiales*. Editorial Centro de Estudios Ramón Areces.

Shakespeare, W. (1989). The King Richard lll. In *The Illustrated Stratford* (pp. 552–614). Chancellor Press.

Wiener, N. (1964). *God and Golem, Inc*. The MIT Press.

CHAPTER 3

The Impact of AI on Business, Economics and Innovation

Ana Landeta Echeberria

Abstract This chapter, based on previous bibliographic review, offers guidance from consultants as well as prior identification of documented professional practices about the current overall impact of Artificial Intelligence (AI) in business, economics and innovation.

Consequently, this chapter discusses the role of AI in the future economy considering increases in productivity, innovation and technological maturity. Consequently, attention shall be paid to business-oriented design, AI tools for business processes modelling and the benefits of AI technologies. It will also be shown how business leaders can remain competitive in the new economic environment, developing the required skills for understanding the economic implications of AI, considering the changes that businesses will need to do to address the economic and social implications of large-scale applications of AI. In addition, we highlight the importance; benefits and applications of Machine Learning in business shall be applied.

A. L. Echeberria (✉)
Faculty of Economics and Business, Universidad a Distancia de Madrid (UDIMA), Madrid, Spain
e-mail: ana.landeta@udima.es

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_3

Finally, the conclusions propose a future research agenda for AI for certain industries (Strategy, Relationship Marketing, Servicescape, Customer acceptance, Social acceptance, Management, Workforce and Transhumanism).

Keywords Business • Economics • Innovation • Artificial Intelligence • Machine learning • Applications • Benefits

3.1 Introduction

In the first place, it must be considered that business Artificial Intelligence (AI) is considered an emerging technology with many future applications, albeit research has been conducted in but a few of them.

Considering this, AI is regarded as one of the most impactful technologies in the word today. This technology powers many real-world applications and also brings both societal benefits and challenges. Businesses and Economies stand to benefit from AI, through increased productivity and innovation.

The impact of AI is compared with the effects the steam engine had on the economy and society in the seventeenth century (Holtel, 2016; Bughin et al., 2018). It is accepted that AI and its applications profoundly impact organizations in various ways (Makridakis, 2017): By implementing the technology into processes and tasks, AI is reshaping jobs, employment and working environments (Huang & Rust, 2018). To understand the effects of AI on organizations, the assumption is that it not primarily aims to replace tasks or even jobs, but that it rather pursues a human-machine cooperation approach (Makridakis, 2017; Hoc, 2000).

This task-oriented approach requires well-documented reference processes and sub-processes (tasks) as well as an understanding of AI which is based on the human perception of intelligence.

According to Dietzmann and Alt (2020) the logic behind the latter is that AI is based on the human experiences with—mostly human—intelligence, because AI is developed by humans. Thus, the evaluation of the impacts of AI on organizations calls for an understanding of both human and AI intelligence capabilities.

3.2 Business Benefits of AI

Many businesses take up AI technology to try to reduce operational costs, increase efficiency, grow revenue and improve customer experience.

This rapidly developing technology offers significant development opportunities that many companies have already been quick to seize upon.

AI systems provide businesses with a wide range of benefits, including personalized marketing, customer service, operational automation, inventory management and recruitment.

Therefore, and in accordance with Web Hosting Blog from eUKhost (2019), some of the ways a company could benefit from AI will be the following ones:

- Improving personalized shopping experiences: Providing customers with personalized marketing increases engagement, helps generate customer loyalty and improves sales. AI is able to provide highly accurate offers to individual customers.
- Automating customer interactions: AI enables companies to automate customer interactions communications by intelligent machines that are able to replicate human functions.
- Real-time assistance: Daily communication with high volumes of customers throughout each day. AI is able to interact, in real-time, to send personalized travel information, such as notice of delays.
- Data mining: One of the biggest advantages of using cloud-based AI is that AI apps are able to quickly discover important and relevant findings during the processing of big data. This can provide businesses with previously undiscovered insights that can help give it an advantage in the marketplace.
- Operational automation: AI is able to operate other technologies that increase automation in business.
- Predicting outcomes: AI is able to predict outcomes based on data analysis.
- Improve the recruitment process: AI is now helping businesses automate the recruitment of new employees. It is able to quickly sift through applications, automatically rejecting those, which do not meet the company's personal specification.

3.2.1 The Business-Oriented Conception of AI Technologies

Therefore, the context of design of clearly business-oriented AI technologies will be subject to the evolution of such technologies, considered in all cases as emerging technologies and, at the same time, innovating instruments with a huge disruption capacity.

Companies around the globe are seeing their industries disrupted by new technologies that result in business model innovation (Wåge & Crawford, 2017). AI—"Intelligent systems created to use data, analysis and observations to perform certain tasks without needing to be programmed to do so" (Antonescu, 2018)—represents the most important technological development. AI disrupts industries and companies when companies use it to create innovative new business models (Boitnott, 2019). Companies such as Amazon, Uber, Tesla, Google, Alibaba and UPS, along with many other companies, have innovated their business models and enhanced their competitive advantages using AI. Top executives need to embrace an entrepreneurial and innovative mindset and instil this mindset using AI throughout their organizations to remain competitive and viable.

The concept of business model innovation has been put to the forefront of the debate of how companies may preserve their market position (Sosna et al., 2010; Wirtz et al., 2010). The present literature of business model innovation mainly focuses on external antecedents, which may pressure companies to engage in business model innovation (Foss & Saebi, 2017). This pressure may also arise through technological disruptions. Researchers argue that the process of business model innovation is prone to being affected by their environment (Chesbrough, 2010).

Lee et al. (2019) go on to suggest that advances in AI technology and data analytics will continue to create opportunities and challenge delivery systems. Effective leaders need to find new and innovative ways to leverage these advances to transform their organization and drive growth. Many leaders will find that these advances take them in directions they had never considered. The focus on a company's core competencies and business strategy remains imperative but remaining open to and encouraging innovations that shift the company's business model represent a major challenge for leaders.

The authors go on to suggest that three preconditions pave the way to success: Transform the core business, grow the core business and scale new business. We believe that for many companies, transforming and growing the core business will take priority over scaling a new business since companies must focus most of their attention on the business generating the most revenues. If companies start new businesses, these will likely be a byproduct of attempts to innovate around legacy businesses or perhaps more likely, legacy systems. Companies are increasingly looking for employees who can innovate and who have an entrepreneurial mindset because they recognize the need for employees who see major problems, view them as opportunities and implement innovative solutions.

However, although the transformative component in the core business and sheer innovation components are very much present, we are obliged to deepen our advances and impact of robotics in economy and current business models, as stated below.

3.2.1.1 AI Impact on Business

Although, digital age together with other sciences like mechatronics, nanotechnology, genetic and so on is a step for "Space Economics,[1]" some other progresses are going to change business and economics directly or indirectly more than other developments (Dirican, 2015). These progresses are named Robotics and AI. The "Industrial Age" has been started by the industrial revolution and mechanization primarily in United Kingdom and by car makers (McKenzie, 2015). Production and deriving from that the supply side of the economics have made huge impacts on business and economics at the beginning of the twentieth century. The Production Factors, that is capital, entrepreneurship, work force and land, were affected by the industrial-age developments and mechanization, and life style, education, finance, and management have been all changed due to these effects (Mokyr, 1985). In order to solve new issues and problems, white collars and management have come to the agenda which created higher education needs due to the level of information, decisions and quality of the work force needed (Keller, 1983). Workers in order to be at the same working time at factories or production lines have begun to live in housing estates, large buildings or complexes which directed life style to live in cities rather than villages. Wages and wealth have changed the buying attitudes and social behaviours (Davies, 1962). Lowering the human work force costs on one side by bulk population management, for

[1] The Space Economy is defined by OECD as the full range of activities and the use of resources that create value and benefits to human beings in the course of exploring, researching, understanding, managing and utilizing space (OECD, 2012).

example transportation, municipal services and employee rights on the business environment, while on the other hand depreciation of allowance of the machines in production lines, and calculation of the return on investment of these production lines led to new definitions in accounting and finance and the cost of capital. Robotics and AI will be also opening new pages in the economics and business which are also bringing new life style and sociological side effects. Roubini and Stiglitz mentioned about the possible results and impacts of these effects in their articles (Roubini, 2014; Stiglitz, 2014), together with many discussions held in the last World Economic Forum 2015 (WEF, 2015) and papers, news are newly started to be issued on the same topics.

By deploying the right AI technology, your business may gain an ability to (Nibusinessinfo, n.d.):

- save time and money by automating and optimizing routine processes and tasks;
- increase productivity and operational efficiencies;
- make faster business decisions based on outputs from cognitive technologies;
- avoid mistakes and "human error", provided that AI systems are set up properly;
- use insight to predict customer preferences and offer them better, personalized experience;
- mine vast amount of data to generate quality leads and grow your customer base;
- increase revenue by identifying and maximizing sales opportunities; and
- grow expertise by enabling analysis and offering intelligent advice and support.

And, according to Infosys (2016), the main driving force for using AI in business was competitor advantage. After that, the incentive came from:

- an executive-led decision,
- a particular business, operational or technical problem,
- an internal experiment,
- customer demand,
- an unexpected solution to a problem and
- an offshoot of another project.

3.2.2 AI for Business Leaders

Rahman (2019) for our purposes, we're going to use one label ("AI"), and define it in business-oriented terms. Business audiences typically prefer an emphasis on actions, outputs and business value. So to frame the discussion for business audiences, here's the definition of AI we use at AI Prescience:

For business, "AI" means computer-based systems that perform, enhance or transform business activities by simulating human behaviour, with the aim of improving business efficiency or effectiveness.

At one level this is a pretty straightforward view of AI, with an emphasis on business results. However, what matters to business is whether AI can achieve better results in areas that used to need human beings.

This perspective treats AI as just another step in the evolution of computerization and even industrialization. What sets this step apart from earlier technologies is that it performs (and increasingly improves) activities we thought only people could do.

It can do this because it consists of computer systems with characteristics we thought were reserved for humans. Examples of human characteristics AI simulates include:

- Awareness and recognition of our surroundings (such as cameras in self-drive cars),
- Interpreting and ascribing "meaning" to information (such as converting hand-drawn sketches and notes into HTML web pages),
- Handling ambiguity (such as knowing whether someone searching online for "golf clubs" wants buy golf equipment or find places to play golf),
- Applying judgement (such as assessing if a strange insurance claim is likely to be fraudulent) and
- Making decisions, even with incomplete information (such as weighing up whether to follow an inconclusive cancer screening with a biopsy, another screening test or an "all clear").

Before AI, if a task required any of these characteristics, we needed people to do it. Now, we have the possibility of automating some or all of it. If the result is better, whether the intelligence is "real" or artificial isn't particularly relevant as far as business results are concerned.

That's not to say business leaders shouldn't care about the other implications of AI, such as ethical dilemmas and impact on the workforce. If anything, the opposite. It's just that if their yardstick for making decisions is solely business results, potentially complex AI choices may become rather straightforward. But that takes us into a very different discussion, one for another time.

For now, the main point for business leaders is that AI technology in business is all about improving activities that previously needed people to do them.

The implication for a business audience is to benefit from AI; businesses have to obtain and manage huge amounts of data, using substantial computer processing, storage and tools to extract value from it. This may need new skills, not previously relevant in most businesses.

Therefore, one of the most important things about AI for business leaders to prioritize is selecting the right kind of problem to solve, and the right kind of business improvement to target. This will largely be a function of the knowledge of their AI teams, along with constraints like timescales and budget.

But ideally, it will be something about which business leaders also have a degree of understanding themselves. This is because applying AI in business is a veritable minefield of unexpected results and unintended consequences. We seem to be seeing increasing numbers of cases where business leaders pay the price for high-profile IT problems, not technology leaders.

Rahman (2019) suggests the following takeaways for Business Leaders.

- To use AI effectively in your business, you need data, in huge amounts and diverse forms.
- You'll also need more computing power and storage than you've probably needed before, and additional skills to those in your current technology and business teams.
- As you start to use AI technology, the more data you get, the more ways you're likely to discover to use it.
- And as the ways you could use data and AI increase, so will the opportunities to make and save money.
- But this will go hand-in-hand with increased possibilities to do inappropriate things with data and AI technology—inadvertently or consciously.
- Successful use of AI technology to improve business is a collective effort, and as a business leader, you're part of a large team of people that can achieve improvements.

- But when AI causes problems, the responsibility may be less widely shared.
- As a business leader, you're part of a much smaller group that will determine how an organization prepares for and respond to the opportunities and risks around AI in business.
- And because regulators and governments typically struggle to keep pace with technology innovations, you may find yourself facing questions, choices and dilemmas with implications far beyond your own business.

3.3 The Importance of Machine Learning in Business

Without a doubt, the importance of Machine Learning (ML) for business is unquestionable, as are their many applications and foreseeable benefits, as seen below.

In the first place is the design of ML in business as a catalyst instrument for enhancing business scalability and improving business operations for companies across the globe. AI tools and numerous ML algorithms have gained tremendous popularity in the business analytics community.

ML is transforming the business world in several ways due to the power of AI to automate the decision-making process.

In particular, there are a number of ways that ML is already making an impact for companies in every industry: Personalizing the customer experience, improving loyalty and retention, enhancing the hiring process, detecting fraud and streamlining IT operations.

The state of ML in business today according to Bean (2018) is that "ML has become more widely adopted by business". This author establishes a difference between AI and Machine Learning; AI is "the broader concept of machines being able to carry out tasks in a way that we would consider smart", while ML is "a current application of AI based around the idea that we should really just be able to give machines access to data and let them learn for themselves". What these approaches have in common is that Machine Learning, deep learning (DL) and AI have all benefited from the advent of Big Data and quantum computing power. Each of these approaches relies upon access to data and powerful computing capacity.

Automating ML: Early adopters of ML are findings ways to automate ML by embedding processes into operational business environments to drive business value. This is enabling more effective and precise learning and decision-making in real-time. When a firm employs automated Machine Learning, these models are then updated without human intervention since they are “constantly learning” based on the very latest data.

Real-Time Decision-Making: For businesses today, growth in data volumes and sources—sensor, speech, images, audio and video—will continue to accelerate as data proliferates. As the volume and speed of data available through digital channels continue to outpace manual decision-making, ML can be used to automate ever-increasing streams of data and enable timely data-driven business decisions. Today, organizations can infuse ML into core business processes that are connected with the firm’s data streams with the objective of improving their decision-making processes through real-time learning.

Businesses that are at the forefront in the application of ML are using approaches such as creating a “workbench” for data science innovation or providing a “governed path to production” which enables “data stream model consumption”. Embedding ML into production processes will help ensure timely and more accurate digital decision-making. Organizations can accelerate the roll-out of these platforms in ways that were not achievable in the past through techniques such as the Analytics Workbench and a Run-Time Decision Framework. These techniques provide data scientists with an environment that enables rapid innovation and helps support increasing analytics workloads, while leveraging the benefits of distributed Big Data platforms and a growing ecosystem of advanced analytics technologies. A “run-time” decision framework provides an efficient path to automate into production ML models that have been developed by data scientists in an analytics workbench.

Driving Business Value: Leaders in ML have been deploying “run-time” decision frameworks for years. What is new today is that technologies have advanced to the point where ML capabilities can be deployed at scale with greater speed and efficiency. These advances are enabling a range of new data science capabilities including the acceptance of real-time decision requests from multiple channels while returning optimized decision results, processing of decision requests in real-time through the execution of business rules, scoring of predictive models and arbitrating among a scored decision set, scaling to support thousands of requests per second, and processing responses from channels that are fed back into models for model recalibration. Firms are deploying run-time decision frameworks with embedded and scalable Machine Learning and are achieving notable results:

- Establishment of a demand funnel of 90 use cases for a next-generation advanced analytics platform.
- Reduction of run time for analytics by more than 90%.
- Support for millions of decision requests per day.
- 46% more leads measured versus a control group.
- $100M+ incremental annual revenue measured versus a control group.

This may sound like data science jargon, but firms are experiencing quantifiable business results. For organizations seeking to compete on data, ML has reached the stage of providing a critical business edge.

3.3.1 Applications of Machine Learning

Finally, we will review ML applications in the context of Fourth Industrial Revolution, when ML becomes popular in various application areas, because of its learning capabilities from the past and making intelligent decisions. In the following, we summarize and discuss ten popular application areas of ML technology (Sarker, 2021).

- *Predictive analytics and intelligent decision-making*: A major application field of ML is intelligent decision-making by data-driven predictive analytics (Cao, 2017; Mahdavinejad et al., 2018). The basis of predictive analytics is capturing and exploiting relationships between explanatory variables and predicted variables from previous events to predict the unknown outcome (Han et al., 2011). For instance, identifying suspects or criminals after a crime has been committed, or detecting credit card fraud as it happens. Another application, where ML algorithms can assist retailers in better understanding consumer preferences and behaviour, better manage inventory, avoiding out-of-stock situations, and optimizing logistics and warehousing in e-commerce. Various ML algorithms such as decision trees, support vector machines and artificial neural networks (Sarker et al., 2019a; Witten et al., 2005) are commonly used in the area. Since accurate predictions provide insight into the unknown, they can improve the decisions of industries, businesses and almost any organization, including government agencies, e-commerce, telecommunications, banking and financial services, healthcare, sales and marketing, transportation, social networking and many others.

- *Cybersecurity and threat intelligence*: Cybersecurity is one of the most essential areas of Industry 4.0. (Ślusarczyk, 2018), which is typically the practice of protecting networks, systems, hardware and data from digital attacks (Ślusarczyk, 2018). ML has become a crucial cybersecurity technology that constantly learns by analysing data to identify patterns, better detect malware in encrypted traffic, find insider threats, predict where bad neighbourhoods are online, keep people safe while browsing or secure data in the cloud by uncovering suspicious activity. For instance, clustering techniques can be used to identify cyber-anomalies, policy violations and so on. To detect various types of cyber-attacks or intrusions ML classification models by taking into account the impact of security features are useful (Sarker et al., 2020b). Various deep learning-based security models can also be used on the large scale of security datasets (Sarker, 2021; Xin et al., 2018). Moreover, security policy rules generated by association rule learning techniques can play a significant role to build a rule-based security system (Sarker et al., 2020a). Thus, we can enable cybersecurity professionals to be more proactive in efficiently preventing threats and cyberattacks.
- Internet of Things (IoT) and smart cities: Internet of Things (IoT) is another essential area of Industry 4.0. (Ślusarczyk, 2018), which turns everyday objects into smart objects by allowing them to transmit data and automate tasks without the need for human interaction. IoT is, therefore, considered to be the big frontier that can enhance almost all activities in our lives, such as smart governance, smart home, education, communication, transportation, retail, agriculture, health care, business and many more (Mahdavinejad et al., 2018). Smart city is one of IoT's core fields of application, using technologies to enhance city services and residents' living experiences (Zanella et al., 2014; Zheng et al., 2015). As ML utilizes experience to recognize trends and create models that help predict future behaviour and events, it has become a crucial technology for IoT applications (Sarker et al., 2021). For example, to predict traffic in smart cities, parking availability prediction, estimate the total usage of energy of the citizens for a particular period, make context-aware and timely decisions for the people and so on are some tasks that can be solved using ML techniques according to the current needs of the people.

- Traffic prediction and transportation: Transportation systems have become a crucial component of every country's economic development. Nonetheless, several cities around the world are experiencing an excessive rise in traffic volume, resulting in serious issues such as delays, traffic congestion, higher fuel prices, increased CO2 pollution, accidents, emergencies and a decline in modern society's quality of life (Guerrero-Ibáñez et al., 2018). Thus, an intelligent transportation system through predicting future traffic is important, which is an indispensable part of a smart city. Accurate traffic prediction based on Machine Learning and deep learning modelling can help to minimize the issues (Boukerche & Wang, 2020; Essien et al., 2019; Essien et al., 2020). For example, based on the travel history and trend of travelling through various routes, ML can assist transportation companies in predicting possible issues that may occur on specific routes and recommending their customers to take a different path. Ultimately, these learning-based data-driven models help improve traffic flow, increase the usage and efficiency of sustainable modes of transportation, and limit real-world disruption by modelling and visualizing future changes.
- Healthcare and COVID-19 pandemic: ML can help to solve diagnostic and prognostic problems in a variety of medical domains, such as disease prediction, medical knowledge extraction, detecting regularities in data and patient management (Fatima & Pasha, 2017; Nilashi et al., 2017; Silahtaroğlu & Yılmaztürk, 2021). Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus, according to the World Health Organization (WHO[2]). Recently, the learning techniques have become popular in the battle against COVID-19 (Kushwaha et al., 2020; Lalmuanawma et al., 2020). For the COVID-19 pandemic, the learning techniques are used to classify patients at high risk, their mortality rate and other anomalies (Kushwaha et al., 2020). It can also be used to better understand the virus's origin, COVID-19 outbreak prediction, as well as for disease diagnosis and treatment (Ardabili et al., 2020; Jamshidi et al., 2020). With the help of Machine Learning, researchers can forecast where and when, the COVID-19 is likely to spread, and notify those regions to match the required arrangements. Deep learning also provides exciting solutions to the problems of medical

[2] World Health Organization: WHO. http://www.who.int/.

image processing and is seen as a crucial technique for potential applications, particularly for COVID-19 pandemic (Alakus & Turkoglu, 2020; Oh et al., 2020; Shorten et al., 2021). Overall, Machine Learning and deep learning techniques can help to fight the COVID-19 virus and the pandemic as well as intelligent clinical decisions making in the domain of healthcare.

- E-commerce and product recommendations: Product recommendation is one of the most well-known and widely used applications of Machine Learning, and it is one of the most prominent features of almost any e-commerce website today. ML technology can assist businesses in analysing their consumers' purchasing histories and making customized product suggestions for their next purchase based on their behaviour and preferences. E-commerce companies, for example, can easily position product suggestions and offers by analysing browsing trends and click-through rates of specific items. Using predictive modelling based on ML techniques, many online retailers, such as Amazon (Marchand & Marx, 2020), can better manage inventory, prevent out-of-stock situations and optimize logistics and warehousing. The future of sales and marketing is the ability to capture, evaluate and use consumer data to provide a customized shopping experience. Furthermore, ML techniques enable companies to create packages and content that is tailored to the needs of their customers, allowing them to maintain existing customers while attracting new ones.
- Natural language processing (NLP) and sentiment analysis: Natural language processing (NLP) involves the reading and understanding of spoken or written language through the medium of a computer (Otter et al., 2020; Sarker et al., 2020c). Thus, NLP helps computers, for instance, to read a text, hear speech, interpret it, analyse sentiment and decide which aspects are significant, where ML techniques can be used. Virtual personal assistant, Chatbot, speech recognition, document description, language or machine translation and so on are some examples of NLP-related tasks. Sentiment analysis (Ravi & Ravi, 2015) (also referred to as opinion mining or emotion AI) is an NLP sub-field that seeks to identify and extract public mood and views within a given text through blogs, reviews, social media, forums, news and so on. For instance, businesses and brands use sentiment analysis to understand the social sentiment of their brand, product or service through social media platforms or the web as a

whole. Overall, sentiment analysis is considered as a ML task that analyses texts for polarity, such as "positive", "negative" or "neutral" along with more intense emotions like very happy, happy, sad, very sad, angry, have interest or not interested.

- Image, speech and pattern recognition: Image recognition (Fujiyoshi et al., 2019) is a well-known and widespread example of ML in the real world, which can identify an object as a digital image. For instance, to label an X-ray as cancerous or not, character recognition, or face detection in an image, and tagging suggestions on social media, for example Facebook, are common examples of image recognition. Speech recognition (Chiu et al., 2018) is also very popular that typically uses sound and linguistic models, for example Google Assistant, Cortana, Siri, Alexa and so on (López et al., 2017), where ML methods are used. Pattern recognition (Anzai, 2012) is defined as the automated recognition of patterns and regularities in data, for example image analysis. Several ML techniques such as classification, feature selection, clustering and sequence labelling methods are used in the area.
- Sustainable agriculture: Agriculture is essential to the survival of all human activities (Sharma et al., 2020). Sustainable agriculture practices help to improve agricultural productivity while also reducing negative impacts on the environment (Adnan et al., 2018; Cobuloglu & Büyüktahtakın, 2015; Sharma et al., 2020). The sustainable agriculture supply chains are knowledge-intensive and based on information, skills, technologies and so on, where knowledge transfer encourages farmers to enhance their decisions to adopt sustainable agriculture practices utilizing the increasing amount of data captured by emerging technologies, for example the Internet of Things (IoT), mobile technologies and devices and so on (Adnan et al., 2018; Kamble et al., 2018; Kamble et al., 2020). ML can be applied in various phases of sustainable agriculture, such as in the pre-production phase—for the prediction of crop yield, soil properties, irrigation requirements and so on; in the production phase—for weather prediction, disease detection, weed detection, soil nutrient management, livestock management and so on; in processing phase—for demand estimation, production planning and so on; and in the distribution phase—the inventory management, consumer analysis and so on.

- *User behaviour analytics and context-aware smartphone applications*: Context-awareness is a system's ability to capture knowledge about its surroundings at any moment and modify behaviours accordingly (Dey, 2001; Sarker, 2019). Context-aware computing uses software and hardware to automatically collect and interpret data for direct responses. The mobile app development environment has been changed greatly with the power of AI, particularly, ML techniques through their learning capabilities from contextual data (Sarker et al., 2020a; Zhu et al., 2012). Thus, the developers of mobile apps can rely on ML to create smart apps that can understand human behaviour, support and entertain users (Sarker & Salah, 2019). To build various personalized data-driven context-aware systems, such as smart interruption management, smart mobile recommendation, context-aware smart searching and decision-making that intelligently assist end mobile phone users in a pervasive computing environment, ML techniques are applicable. For example, context-aware association rules can be used to build an intelligent phone call application (Sarker & Kayes, 2020). Clustering approaches are useful in capturing users' diverse behavioural activities by taking into account data in time series (Sarker et al., 2018). To predict the future events in various contexts, the classification methods can be used (Sarker et al., 2019a). Thus, various learning techniques discussed in Sect. "Machine Learning Tasks and Algorithms[3]" can help to build context-aware adaptive and smart applications according to the preferences of the mobile phone users.

In addition to these application areas, Machine Learning-based models can also apply to several other domains such as bioinformatics, cheminformatics, computer networks, DNA sequence classification, economics and banking, robotics, advanced engineering, and many more (Sarker, 2021).

[3] ML algorithms that include classification analysis, regression analysis, data clustering, association rule learning, feature engineering for dimensionality reduction, as well as deep learning methods. See at https://link.springer.com/content/pdf/10.1007/s42979-021-00592-x.pdf.

3.3.2 Business Benefits of Machine Learning

Machine Learning in business helps in enhancing business scalability and improving business operations for companies across the globe. AI tools and numerous ML algorithms have gained tremendous popularity in the business analytics community. Factors such as growing volumes, easy availability of data, cheaper and faster computational processing and affordable data storage have led to a massive ML boom. Therefore, organizations can now benefit by understanding how businesses can use ML and implement the same in their own processes, as Flatworld Solutions Pvt. Ltd. stated.

It is because of that ML helps in extracting meaningful information from a huge set of raw data. If implemented in the right manner, ML can serve as a solution to a variety of business complexities problems and predict complex customer behaviours. We have also seen some of the major technology giants, such as Google, Amazon and Microsoft, coming up with their Cloud ML platforms.

Therefore, some of the key ways in which ML can help your business according to the same consultant are the following:

1. Customer Lifetime Value Prediction: Customer lifetime value prediction and customer segmentation are some of the major challenges faced by the marketers today. Companies have access to huge amount of data, which can be effectively used to derive meaningful business insights. ML and data mining can help businesses predict customer behaviours, purchasing patterns, and help in sending best possible offers to individual customers, based on their browsing and purchase histories.
2. Predictive Maintenance: Manufacturing firms regularly follow preventive and corrective maintenance practices, which are often expensive and inefficient. However, with the advent of ML, companies in this sector can make use of ML to discover meaningful insights and patterns hidden in their factory data. This is known as predictive maintenance and it helps in reducing the risks associated with unexpected failures and eliminates unnecessary expenses. ML architecture can be built using historical data, workflow visualization tool, flexible analysis environment and the feedback loop.
3. Eliminates Manual Data Entry: Duplicate and inaccurate data are some of the biggest problems faced by the businesses today. Predictive modelling algorithms and ML can significantly avoid

any errors caused by manual data entry. ML programs make these processes better by using the discovered data. Therefore, the employees can utilize the same time for carrying out tasks that add value to the business.

4. Detecting Spam: ML in detecting spam has been in use for quite some time. Previously, email service providers made use of pre-existing, rule-based techniques to filter out spam. However, spam filters are now creating new rules by using neural networks detect spam and phishing messages.
5. Product Recommendations: Unsupervised learning helps in developing product-based recommendation systems. Most of the e-commerce websites today are making use of ML for making product recommendations. Here, the ML algorithms use customer's purchase history and match it with the large product inventory to identify hidden patterns and group similar products together. These products are then suggested to customers, thereby motivating product purchase.
6. Financial Analysis: With large volumes of quantitative and accurate historical data, ML can now be used in financial analysis. ML is already being used in finance for portfolio management, algorithmic trading, loan underwriting and fraud detection. However, future applications of ML in finance will include Chatbots and other conversational interfaces for security, customer service and sentiment analysis.
7. Image Recognition: Also known as computer vision, image recognition has the capability to produce numeric and symbolic information from images and other high-dimensional data. It involves data mining, ML, pattern recognition and database knowledge discovery. ML in image recognition is an important aspect and is used by companies in different industries including healthcare, automobiles and so on.
8. Medical Diagnosis: ML in medical diagnosis has helped several healthcare organizations to improve the patient's health and reduce healthcare costs, using superior diagnostic tools and effective treatment plans. It is now used in healthcare to make almost perfect diagnosis, predict readmissions, recommend medicines and identify high-risk patients. These predictions and insights are drawn using patient records and datasets along with the symptoms exhibited by the patient.

9. Improving Cyber Security: ML can be used to increase the security of an organization, as cyber security is one of the major problems solved by Machine Learning. Here, ML allows new-generation providers to build newer technologies, which quickly and effectively detect unknown threats.
10. Increasing Customer Satisfaction: ML can help in improving customer loyalty and also ensure superior customer experience. This is achieved by using the previous call records for analysing the customer behaviour and based on that the client requirement will be correctly assigned to the most suitable customer service executive. This drastically reduces the cost and the amount of time invested in managing customer relationship. For this reason, major organizations use predictive algorithms to provide their customers with suggestions of products they enjoy.

3.4 The Economic and Social Implications of Large-Scale Applications of AI

The analysis[4] carried out by (Soni et al., 2020) of the top 200 AI start-ups explicitly shows the influence of advance research and innovation in AI on the global market. The study shows that the AI wave is on and an appetite for AI growth is exponential. The investment in AI is showing an upward trajectory in the last six years and should remain the same for the upcoming years. The study also uncovers the top AI industries that will generate more opportunities in near future, viz. business intelligence, healthcare, core AI, cybersecurity, and marketing and sales. Some of the key advantages of automation, cognitive technologies and data analysis using AI algorithms are an increase in productivity, time and cost efficiency, human error reduction, faster business decisions, customer preference prediction and sales maximization. However, the study shows that the AI technology is confined only in a few regions in the world. This is creating an "AI divide". This divide, like the digital divide, would strengthen the inequality in social, economic and cultural sectors; would create a chasm. Moreover, AI is software dominant and software is prone to vulnerabilities. Some of the deep learning algorithms/methods are the backbone of AI; these require passing through multiple factors to be used for real-time applications. Identifiable systemic failure modes, repeatability,

[4] AI in business: From research and innovation development to market.

transparency, explainable, path tracing, penetrability and so on are some of the major factors established at the time of assessment of software; even after passing through these factors, there exist cases where DL algorithms have produced unreliable results. Apart from these, challenges like trust, ethics, bias and shortage of AI talent also need attention for commercial usage of AI applications.

Economists are generally enthusiastic about the prospects of AI on economic growth. Economic literature has linked innovation to economic growth (Romer, 1990).

3.4.1 AI and the Economy

Furman and Seamans (2019) reviewed the evidence that AI is having a large effect on the economy.

Across a variety of statistics—including robotics shipments, AI start-ups, and patent counts—there is evidence of a large increase in AI-related activity. We also review recent research in this area that suggests that AI and robotics have the potential to increase productivity growth but may have mixed effects on labour, particularly in the short run. In particular, some occupations and industries may do well while others experience labour market upheaval. We then consider current and potential policies around AI that may help to boost productivity growth while also mitigating any labour market downsides, including evaluating the pros and cons of an AI-specific regulator, expanded antitrust enforcement, and alternative strategies for dealing with the labour market impacts of AI, including universal basic income and guaranteed employment.

AI has been advancing rapidly in recent years, measured both in terms of the amount of resources devoted to it and also in terms of its outputs.[5]

3.4.2 Economic Characteristics of AI

In accordance with OECD (2019), from an economic point of view, recent advances in AI either decrease the cost of prediction or improve the quality of predictions available at the same cost. Many aspects of decision-making are separate from prediction. However, improved, inexpensive

[5] Artificial Intelligence is a loose term used to describe a range of advanced technologies that exhibit human-like intelligence including machine learning, autonomous robotics and vehicles, computer vision, language processing, virtual agents and neural networks.

and widely accessible AI prediction could be transformative because prediction is an input into much of human activity. As the cost of AI prediction has decreased, more opportunities to use prediction have emerged, as with computers in the past. The first AI applications were long-recognized as prediction problems. For example, ML predicts loan defaults and insurance risk. As their cost decreases, some human activities are being reframed as prediction issues. In medical diagnosis, for example, a doctor uses data about a patient's symptoms and fills in missing information about the cause of those symptoms. The process of using data to complete missing information is a prediction. Object classification is also a prediction issue: A human's eyes take in data in the form of light signals and the brain fills in the missing information of a label. AI, through less expensive prediction, has a large number of applications because prediction is a key input into decision-making. In other words, prediction helps make decisions, and decision-making is everywhere. Managers make important decisions around hiring, investments and strategy, and less important decisions around which meetings to attend and what to say during these meetings. Judges make important decisions about guilt or innocence, procedures and sentencing, and smaller decisions about a specific paragraph or motion. Similarly, individuals make decisions constantly—from whether to marry to what to eat or what song to play. A key challenge in decision-making is dealing with uncertainty. Because prediction reduces uncertainty, it is an input into all these decisions and can lead to new opportunities.

Implementing AI in organizations requires complementary investments and process changes. Like computing, electrification and the steam engine, AI can be seen as a general-purpose technology (Bresnahan & Trajtenberg, 1992; Brynjolfsson et al., 2017). This means it has potential to substantially increase productivity in a wider variety of sectors. At the same time, the effect of AI requires investment in a number of complementary inputs. It may lead an organization to change its overall strategy. In the AI context, organizations need to make a number of complementary investments before AI has a significant impact on productivity. These investments involve infrastructure for the continued collection of data, specialized workers that know how to use data and changes in processes that take advantage of new opportunities arising from reduced uncertainty. Many processes in every organization exist to make the best of a situation in the face of uncertainty rather than to serve customers in the best way possible. Airport lounges, for example, make customers comfortable while they wait for their plane. If passengers had accurate predictions of how long it would take to get to the airport and through security, lounges might not be needed. The scope of opportunities offered by better predictions is expected to vary across companies and industries. Google, Baidu and other large digital platform companies are well-positioned to benefit from major investments in AI. On the supply side, they already have systems in place to collect data. On the demand side, having enough customers to justify the high fixed costs of investment in the technology is in its early stages. Many other businesses have not fully digitized their workflows and cannot yet apply AI tools directly into existing processes. As costs fall over time, however, these businesses will recognize the opportunities that are possible by reducing uncertainty. Driven by their needs, they will follow industry leaders and invest in AI.

3.4.3 AI in Business: Future Research Agenda

In order to guide the directions of future research addressing AI in business, Loureiro et al. (2020) propose a set of research questions for certain business domains (Strategy, Relationship Marketing, Servicescape, Customer acceptance, Social acceptance, Management, Workforce and Transhumanism) that are still in need for further development.

Business Domain: Strategy
Research Questions:

- How to design human-machine integrated service strategies?
- How to create original, unique goods with AI applications?
- How to identify consumer preferences for human or machine services?
- How to streamline processes for human and machine service providers?
- How hard is to duplicate competitive advantages based on robots and automated systems?
- How robots and automated systems depreciate over time and what kind of investment is required to keep the pace of innovation under an AI-led business environment?

Business Domain: Relationship Marketing
Research Questions:

- How will cognitive and emotional-social complexity dimensions influence robot design?

- Which relationship marketing capabilities should be programmed in service robot?
- How will the engagement process evolve between humans and AI-enabled machines? And, what about enabled-AI Robot-to-Robot engagement?
- How can robots and automated systems based on AI empower consumers with disabilities to go beyond their cognitive and physical limitations?

Business Domain: Servicescape
Research Questions:

- How can service robots be effectively integrated into the servicescape?
- How will the new servicescape look like in a robot dominated service environment?
- What will be the core dimensions of servicescape in the service robots?

Business Domain: Customer acceptance
Research Questions:

- Beyond the physical and virtual nature of service objects, what drives customer preference for physical or virtual robots?
- How can the more cost-effective virtual robots (e.g. holograms rather than physical robots at information counters) be designed to achieve greater consumer acceptance?
- Which consumer and context factors determine the optimal level of humanoid appearance and social skills for service robots?
- How robot gender and personality will impact consumer responses to service robots?
- Which service and industry characteristics will potentially moderate the impacts of determinants of customer acceptance of service robots?

Business Domain: Social acceptance
Research Questions:

- How can employees deal with working alongside humanoid robots that can operate for extended periods without human intervention (making their own decisions and acting independently)?

- How can humans deal with autonomous systems in social environments?
- Can humanoid robots be accepted to be part of social events (professional or not) as any human?

Business Domain: Management
Research Questions:

- How should organizations manage and implement AI systems in their organizations?
- How should organizations stimulate employees to use such systems?
- How can we re-train workers for intuitive and empathetic skills to remain employable?
- How can we educate students for intuitive and empathetic skills to remain employable in an AI-led business environment?
- How can organizations improve knowledge creation with automation using generative models of deep learning?

Business Domain: Workforce
Research Questions:

- How can we engage stakeholders to develop new product/experiences embracing Brain-Computer Interface (BCI) and Direct Broadcast Service (DBS) technologies?
- How can we engage employees in the application of transhumanistic technologies to improve job performance?
- What is the impact for the company of recruiting employees with BCI and DBS technologies?

Business Domain: Transhumanism
Research Questions:

- How to connect humans and/or AI-enabled machines for collective intelligence?
- How may transhumanistic technologies enhance human well-being?
- How citizens/consumers/employees who decide not to use BCI and DBS technologies will interact with those who do?

Despite using prior research as a basis for proposing future directions, the suggested questions may still be far from being fully answered as no one knows the pace of AI evolution.

References

Adnan, N., Shahrina, M. N., Rahman, I., & Noor, A. (2018). The effects of knowledge transfer on farmers decision making toward sustainable agriculture practices. *World Journal of Science, Technology and Sustainable Development*.

Alakus, T. B., & Turkoglu, I. (2020). Comparison of deep learning approaches to predict COVID-19 infection. *Chaos, Solitons & Fractals, 140*, 110120.

Antonescu, D. (2018). Are business leaders prepared to handle the upcoming revolution in business artificial intelligence? *Quality-Access to Success, 19*(166).

Anzai, Y. (2012). *Pattern recognition and machine learning*. Elsevier.

Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R., Reuter, U., … Atkinson, P. M. (2020). Covid-19 outbreak prediction with machine learning. *Algorithms, 13*(10), 249.

Bean, R. (2018). The state of ML in business today. *Preuzeto od*. https://www.forbes.com/sites/ciocentral/2018/09/17/the-state-of-machine-learning-inbusiness-today

Boitnott, J. (2019). *Innovative companies using AI to distrust their industries.* Retrieved June 1, 2019, from https://www.inc.com/john-boitnott/7-innovative-companies-using-ai-to-disrupt-their-industries.html

Boukerche, A., & Wang, J. (2020). Machine learning-based traffic prediction models for Intelligent Transportation Systems. *Computer Networks, 181*, 107530.

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). *Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics.* Cambridge University Press.

Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Notes from the AI frontier: Modeling the impact of AI on the world economy*. McKinsey Global Institute.

Cao, L. (2017). Data science: A comprehensive overview. *ACM Computing Surveys (CSUR), 50*(3), 1–42.

Chesbrough, H. (2010). Business model innovation: Opportunities and barriers. *Long Range Planning, 43*(2–3), 354–363.

Chiu, C.-C., Sainath, T. N., Wu, Y., Prabhavalkar, R., Nguyen, P., Chen, Z., Kannan, A., Weiss, R. J., Rao, K., Gonina, E., et al. (2018). State-of-the-art speech recognition with sequence-to-sequence models. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4774–4778). IEEE.

Cobuloglu, H., & Büyüktahtakın, I. E. (2015). A stochastic multi-criteria decision analysis for sustainable biomass crop selection. *Expert Systems with Applications, 42*(15–16), 6065–6074.

Davies, J. C. (1962). Toward a theory of revolution. *American Sociological Review*, 5–19.

Dey, A. K. (2001). Understanding and using context. *Personal and Ubiquitous Computing, 5*(1), 4–7.

Dietzmann, C., & Alt, R. (2020, January). Assessing the business impact of Artificial Intelligence. In *Proceedings of the 53rd Hawaii International Conference on System Sciences.*

Dirican, C. (2015). The impacts of robotics, artificial intelligence on business and economics. *Procedia-Social and Behavioral Sciences, 195*, 564–573.

Essien, A., Petrounias, I., Sampaio, P., & Sampaio, S. (2019). Improving urban traffic speed prediction using data source fusion and deep learning. In *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 1–8). IEEE.

Essien, A., Petrounias, I., Sampaio, P., & Sampaio, S. (2020). A deep-learning model for urban traffic flow prediction with traffic events mined from twitter. *World Wide Web*, 1–24.

Fatima, M., & Pasha, M. (2017). Survey of ML algorithms for disease diagnostic. *Journal of Intelligent Learning Systems and Applications, 9*(1), 1.

Flatworld Solutions Pvt. (Ltd). (n.d.). Top 10 ways in which machine learning can help your business. https://www.flatworldsolutions.com/IT-services/articles/how-machine-learning-can-help-your-business.php

Foss, N. J., & Saebi, T. (2017). Fifteen years of research on business model innovation: How far have we come, and where should we go? *Journal of Management, 43*(1), 200–227.

Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Deep learning-based image recognition for autonomous driving. *IATSS Research, 43*(4), 244–252.

Furman, J., & Seamans, R. (2019). AI and the economy. *Innovation Policy and the Economy, 19*(1), 161–191.

Guerrero-Ibáñez, J., Zeadally, S., & Contreras-Castillo, J. (2018). Sensor technologies for intelligent transportation systems. *Sensors, 18*(4), 1212.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques*. Elsevier.

Hoc, J. M. (2000). From human–machine interaction to human–machine cooperation. *Ergonomics, 43*(7), 833–843.

Holtel, S. (2016). Artificial intelligence creates a wicked problem for the enterprise. *Procedia Computer Science, 99*, 171–180.

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155–172.

Infosys (Firm). (2016). Amplifying human potential: Education and skills for the fourth industrial revolution.

Jamshidi, M., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., … Mohyuddin, W. (2020). Artificial intelligence and COVID-19: Deep learning approaches for diagnosis and treatment. *IEEE Access, 8*, 109581–109595.

Kamble, S. S., Gunasekaran, A., & Gawankar, S. A. (2018). Sustainable industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. *Process Safety and Environment Protection, 117*, 408–425.

Kamble, S. S., Gunasekaran, A., & Gawankar, S. A. (2020). Achieving sustainable performance in a data-driven agriculture supply chain: A review for research and applications. *International Journal of Production Economics, 219*, 179–194.

Keller, G. (1983). *Academic strategy: The management revolution in American higher education*. JHU Press.

Kushwaha, S., Bahl, S., Bagha, A. K., Parmar, K. S., Javaid, M., Haleem, A., & Singh, R. P. (2020). Significant applications of ML for COVID-19 pandemic. *Journal of Industrial Integration and Management, 5*(4).

Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020). Applications of ML and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons & Fractals, 139*, 110059.

Lee, J., Suh, T., Roy, D., & Baucus, M. (2019). Emerging technology and business model innovation: The case of artificial intelligence. *Journal of Open Innovation: Technology, Market, and Complexity*, 5, 44–3.

López, G., Quesada, L., & Guerrero, L. A. (2017). Alexa vs. Siri vs. Cortana vs. Google assistant: A comparison of speech-based natural user interfaces. In *International Conference on Applied Human Factors and Ergonomics* (pp. 241–250). Springer.

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2020). Artificial intelligence in business: State of the art and future research agenda. *Journal of Business Research*.

Mahdavinejad, M. S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., & Sheth, A. P. (2018). Machine Learning for Internet of Things data analysis: A survey. *Digital Communications and Networks, 4*(3), 161–175.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures, 90*, 46–60.

Marchand, A., & Marx, P. (2020). Automated product recommendations with preference-based explanations. *Journal of Retailing, 96*(3), 328–343.

McKenzie, S. (2015). Rise of the robots: The evolution of Ford's assembly line. *CNN Money News*. http://money.cnn.com/gallery/technology/2015/04/29/ford-factory-assembly-line-robots/1.html

Mokyr, J. (1985). *The economics of the industrial revolution* (pp. 1–30). Rowman & Littlefield Publishers, Inc. (Reprinted 1989).

Nibusinessinfo. (n.d.). *Artificial intelligence in business*. https://www.nibusinessinfo.co.uk/content/business-benefits-artificial-intelligence

Nilashi, M., bin Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using ML techniques. *Computers & Chemical Engineering, 106*, 212–223.

OECD. (2019). *Artificial intelligence in society*. OECD Publishing. https://doi.org/10.1787/eedfee77-en

OECD. Publishing, & Organisation for Economic Co-operation and Development Staff. (2012). *OECD handbook on measuring the space economy*. OECD Publishing.

Oh, Y., Park, S., & Ye, J. C. (2020). Deep learning COVID-19 features on CXR using limited training data sets. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2020.2993291

Otter, D. W., Medina, J. R., & Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. *EEE Transactions on Neural Networks and Learning Systems*.

Rahman, W. (2019, December 3). An introduction to artificial intelligence for business leaders. https://towardsdatascience.com/https-medium-com-aiprescience-an-introduction-to-artificial-intelligence-for-business-leaders-93b6fe3d2163

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-based Systems, 89*, 14–46.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5), S71–S102.

Roubini, N. (2014). Rise of the machines: Downfall of the economy. *Nouriel Unplugged: Economic Insights of a Global Nomad*.

Sarker, I. H. (2019). Context-aware rule learning from smartphone data: Survey, challenges and future directions. *Journal of Big Data, 6*(1), 1–25.

Sarker, I. H. (2021). Deep cybersecurity: A comprehensive overview from neural network and deep learning perspective. *SN Computer Science, 2*(3), 1–16.

Sarker, I. H., Abushark, Y. B., Alsolami, F., & Khan, A. I. (2020b). Intrudtree: A ML based cyber security intrusion detection model. *Symmetry, 12*(5), 754.

Sarker, I. H., Colman, A., Kabir, M. A., & Han, J. (2018). Individualized time series segmentation for mining mobile phone user behavior. *The Computer Journal, 61*(3), 349–368.

Sarker, I. H., Hoque, M. M., Uddin, M. K., & Alsanoosy, T. (2021). Mobile data science and intelligent apps: Concepts, AI-based modeling and research directions. *Mobile Networks and Applications, 26*(1), 285–303.

Sarker, I. H., Hoque, M. M., Uddin, M. K., & Tawfeeq, A. (2020c). Mobile data science and intelligent apps: concepts, AI-based modeling and research directions. *Mobile Networks and Applications*.

Sarker, I. H., & Kayes, A. S. M. (2020). ABC-ruleminer: User behavioral rule based ML method for context-aware intelligent services. *Journal of Network and Computer Applications*.

Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020a). Cybersecurity data science: An overview from ML perspective. *Journal of Big Data, 7*(1), 1–29.

Sarker, I. H., Kayes, A. S. M., & Watters, P. (2019a). Effectiveness analysis of ML classification models for predicting personalized context-aware smartphone usage. *Journal of Big Data, 6*(1), 1–28.

Sarker, I. H., & Salah, K. (2019). AppsPred: predicting context-aware smartphone apps using random forest learning. *Internet Things., 8.*

Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., & Kumar, A. (2020). A systematic literature review on ML applications for sustainable agriculture supply chain performance. *Computers and Operations Research, 119.*

Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Deep learning applications for covid-19. *Journal of Big Data, 8*(1), 1–54.

Silahtaroğlu, G., & Yılmaztürk, N. (2021). Data analysis in health and big data: AML medical diagnosis model based on patients' complaints. *Communications in Statistics-Theory and Methods, 50*(7), 1547–1556.

Ślusarczyk, B. (2018). Industry 4.0: Are we ready? *Polish Journal of Management Studies, 17.*

Soni, N., Sharma, E. K., Singh, N., & Kapoor, A. (2020). Artificial intelligence in business: From research and innovation to market deployment. *Procedia Computer Science, 167*, 2200–2210.

Sosna, M., Trevinyo-Rodríguez, R. N., & Velamuri, S. R. (2010). Business model innovation through trial-and-error learning: The Naturhouse case. *Long Range Planning, 43*(2–3), 383–407.

Stiglitz, J. E. (2014). *Unemployment and innovation* (No. w20670). National Bureau of Economic Research.

Trajtenberg, M., & Bresnahan, T. F. (1992). *General Purpose Technologies: Engines of Growth* (No. 16–92).

Wåge, D., & Crawford, G. E. (2017). *Innovation in digital business models.* In Proceedings of the 2017 XXVIII ISPIM Innovation Conference on Composing the Innovation Symphony, Vienna, Austria, 18–21 June 2017.

Web Hosting Blog from eUKhost. (2019). 7 benefits of artificial intelligence for business. Retrieved July 1, 2021, from https://www.eukhost.com/blog/webhosting/7-benefits-of-artificial-intelligence-for-business/

WEF. World Economic Forum Annual Meeting 2015 Programme. (21–24 January 2015). Davos-Klosters, Switzerland. http://www3.weforum.org/docs/AM15/WEF_AM15_Programme_A3526EC2F4.pdf

Wirtz, B. W., Schilke, O., & Ullrich, S. (2010). Strategic development of business models: implications of the Web 2.0 for creating value on the internet. *Long Range Planning, 43*(2–3), 272–290.

Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., & DATA, M. (2005). Practical ML tools and techniques. In *Data Mining* (vol. 2, p. 4).

Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., … Wang, C. (2018). ML and deep learning methods for cybersecurity. *IEEE Access, 6*, 35365–35381.
Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things Journal, 1*(1), 22–32.
Zheng, Y., Rajasegarar, S., & Leckie, C. (2015). Parking availability prediction for sensor-enabled car parks in smart cities. In *2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)* (pp. 1–6). IEEE.
Zhu, H., Cao, H., Chen, E., Xiong, H., & Tian, J. (2012). Exploiting enriched contextual information for mobile app classification. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 1617–1621). ACM.

CHAPTER 4

AI Implications for the Future of Work

María Aránzazu de las Heras García

Abstract Over the course of history, people have tried to find ways to improve their quality of life and to make their jobs easier, and that has required innovation and, more recently, technology.

In barely three decades, the appearance of the Internet, social networks, mobile phones, big data, cloud computing, cybersecurity, robotics or Artificial Intelligence (AI) have wrought enormous, far-reaching changes in the world of work. Their appearance has brought changes in production, in the economy and finances, in cities, trade, health care, agriculture, education and so on. All of that has altered the way we work, consume and communicate with each other, moving from highly personal relations to relations between people and machines. The technological change we are experiencing presages a disruptive transformation in the forms and understanding of work in the future.

Keywords Artificial Intelligence • Jobs • Education • Robots • Automation • Work • Digital transformation

M. A. de las Heras García (✉)
Universidad a Distancia de Madrid (UDIMA), Madrid, Spain
e-mail: arancha.heras@udima.es

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_4

4.1 Key Concepts in Understanding the Implications of AI on the Work of the Future

Before embarking on the study of how Artificial Intelligence (AI) influences the future of the labour market, and by extension Education, it seems necessary to define its key concepts, especially how the digital transformation and AI influence the labour market. Thus:

4.1.1 Digital Transformation in the Workplace

Digitalisation and the digital transformation are the two key terms to understand the changes and the impact, which the evolution of technology is having on society in general.

The digital transformation of our work environment has undergone constant evolution in recent years. For some time now, digitalisation, flexible work schedules, the work-life balance and the new work models have been the main drivers of the digital transformation in work.

The bases of the functioning of the digital transformation are the existence and development of cloud computing, big data, mobile applications, geolocation, the Internet of Things (IoT) and mobile robots, among others (Valenduc & Vendramin, 2016).

The digital transformation has revolutionised the way in which we work, permitting the appearance of new ways of organising work and production, such as teleworking, the delocation of functions, cross-border outsourcing and subcontracting, worldwide supply chains, the open outsourcing of tasks based on the Internet or crowdworking, among others (Cedrola Spremolla, 2019).

That gives rise to important challenges for the sphere of work and human resources, such as the ways of controlling remote work or not violating the right to privacy, the differing ICT skills of different generations, the use of social networks at work, data protection or digital disconnection and so on, because it produces changes not only in the ways of working but also in the production processes.

Therefore, the digital transformation needs to be concerned not only with how to do things but also the processes around it: the evaluation of jobs, the indicators of productivity or motivation, while at the same time managing a considerable degree of resistance to change by all or part of the organisation; the lack of executive competencies; the perception of the high cost of change—financial, psychological and so on; the uncertainty caused by legislation itself, which, on occasions, lags behind the changes in the environment, or the lack of vision on the part of company management (Pérez-González et al., 2018).

Likewise, in this context of change, new leaderships emerge, which must adapt to the new models to direct face-to-face and remote teams, keep people connected and committed, and promote a culture of innovation, learning and constant improvement (Bersin et al., 2017).

4.1.2 Algorithms or Facilitators of Tasks

Algorithms facilitate our daily tasks, and in many cases it is almost impossible to imagine how we could get by without them.

Algorithms are advanced computer programs, which, in companies, by gathering workers' data, make it possible to manage an intrinsic and increasingly important part of the labour relationship by means of automated decision-making.

Algorithms are absolutely not infallible, and their decisions may be just as biased as those of any human being. That was stated by the Council of Europe, in its Recommendation CM/Rec (2020) on the human rights impacts of algorithmic systems, when it affirms that "it is important to note that most algorithmic systems are based on statistical models in which errors form an inevitable part, sometimes with feedback loops that maintain, replicate and reinforce pre-existing biases, errors and assumptions".

Some examples of work-related algorithms can be found in staff recruitment, the organisation of work schedules, the monitoring and tracking of activity, professional promotion, the calculation of performance, the application of the disciplinary system, the determination of salaries or even the evaluation of a dismissal.

4.1.3 Artificial Intelligence (AI) for Work-Related Decision-Making

AI is nothing new. Scientists of the stature of Alan Turing, Marvin Minsky or John McCarthy developed many of its theoretical and technological foundations over the last 70 years.

As indicated by the Commission Communication 2018 to the European Council, the Council of Europe, the European Economic and Social Committee and the Committee of the Regions, entitled AI for Europe[1]:

[1] European Commission, Brussels, 25 April 2018, COM (2018) 237 final.

"Artificial Intelligence (AI) is not science fiction; it is already part of our everyday lives", as also demonstrated by the McKinsey study, which indicates that up to 10% of companies are already applying AI to Human Resources[2] or the Eurofound study, which reports that over 40% of Spanish companies are using big data to monitor their employees.[3]

As well as making our lives easier, AI is helping us to resolve some of the main challenges faced by our world, from the treatment of chronic diseases (e.g. by increasing the precision of diagnoses and permitting better disease prevention, and developing treatments for diseases which previously did not exist) to helping in the fight against climate change (it can help improve weather forecasts and anticipate disasters, it can improve the efficiency of the production systems through predictive maintenance, increase safety, and contribute many other changes which, right now, we can only imagine).

To respond in a similar way to humans, AI must be underpinned by technology, by means of the creation of patterns, which resolve problems and adapt to different situations. That learning is achieved through the gathering of large quantities of data, processing and analysing those data, and thus generating the algorithms or guidelines **for action.**

For Hislop et al. (2017), AI is defined as the development of computers to participate in thought processes similar to those of humans, such as *deep learning*, reasoning and self-correction.

According to the PWC report on essential technologies,[4] the term "AI" refers to software algorithms, which automate complex decision-making tasks, which imitate the senses and processes of human thought. That automatic learning, or "machine learning", focuses on developing computer programs which can teach themselves to learn, reason, plan and act when they receive large quantities of data.

AI is the ability of machines, by means of algorithms, to learn from the data and use them in decision-making, responding in a way, which is similar to the way a human being would do it.

[2] *The state of AI in 2020*, www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020

[3] www.eurofound.europa.eu/publications/report/2020/employee-monitoring-and-surveillance-the-challenges-of-digitalisation

[4] PWC, *The Essential Eight*, 2018 [online]. Available at: https://www.pwc.com/gx/en/issues/technology/essential-eight-technologies.html#content-free-1-f1hc-par

Its specificity lies in the ability to carry out a series of "complex tasks" autonomously, with very little human intervention, and also the ability to adapt to new and changing circumstances, that is to learn based on prior experience (Schlogl & Sumner, 2018).

4.2 AI in the World of Work

AI is now the main driver of economic and productive growth, and it will contribute to the sustainability and viability of the industrial base.[5]

AI will cause changes in our way of working, increasing collaborative work and flexibility in working hours.

The so-called Fourth Industrial Revolution is a phenomenon, which is impacting the labour market, to the extent that it is driving increasing automation of work, which will affect all activities without exception, regardless of salaries or professional qualifications.

The use of technology brings with it the use of robots and AI in different spheres of the production of goods and the provision of services, which is expanding the automation of work, which requires a reassessment of professional expectations for the future, because the owners of capital can gain more by investing in robots which improve the work conditions of their employees (Aguilera Durán, 2019).

The incorporation of robotics and AI into work makes it possible to replace people with robots in many tasks, with greater speed and efficiency. In addition, as stated by Salazar (2019, p. 57), "they do not complain, they do not want days off, they do not stop to eat or sleep, they do not demonstrate against falling wages because they are not paid wages". That is why business owners see these technologies as an opportunity which they must not let escape.

But their use does not bring only advantages, and we are already starting to see some potential risks which must also be taken into account, such as the opacity of decision-making, discrimination among workers, intrusion into our private lives, or the possibility of using them for criminal purposes.

[5] The High-Level Strategy Group on Industrial Technologies recommended that AI be included as one of the key enabling technologies due to its intersectoral potential, which is crucial for European industry. https://publications.europa.eu/en/publication-detail/-/publication/28e1c485-476a-11e8-be1d-01aa75ed71a1/language-en

As outlined in the White Paper on Artificial Intelligence,[6] AI can also cause damage, both tangible (for the health and safety of people) and intangible (loss of privacy, limitations on the right to freedom of expression, human dignity, discrimination in access to employment, etc.).

AI also presents different limitations, the most significant being its own technical limitations, together with the possible reticence of consumers to their use (including the safety problems created by robots inside factories), the legal vacuums that still exist and which stand in the way of their implementation, or the moral debates around the different responsibilities with regard to possible failures of the systems.

AI can also lead to an increase in the precariousness of certain jobs, by limiting their tasks, making it necessary to implement codes of conduct for the protection of human beings from robots.

Therefore, it appears to be necessary to establish a regulatory framework which seeks to minimise the different risks of suffering damage which may arise, because it is undeniable that AI is here to stay, the same as other technological developments which today form part of our daily lives.

That regulatory framework must be flexible, to rapidly accommodate the changes that occur, but also reflexive and limiting, especially with regard to discrimination.

4.2.1 The Impact of Robotisation Processes on the Labour Market

One of the most complex questions, which must be addressed with regard to the future of work, is how the robotisation processes will impact the labour market, in labour markets, which are already weakened, and with high levels of unemployment and job insecurity.

Though there is no consensus on the effects it will have, it is indisputable that its impact will be very considerable and we must remain alert to its evolution.

In general, the studies analysed coincide in indicating that the development of AI and robotics will have consequences, both direct and indirect, especially with regard to the contents of jobs, which will tend to become less routine, while requiring more specialised technological knowledge.

[6] White Paper on Artificial Intelligence; a European approach to excellence and trust, (2020) Brussels, 19.2.2020 COM (2020) 65 final.

According to data from the European Economic and Social Committee (EESC), between 9% and 54% of jobs are under threat, depending on the different methodological options,[7] while the European Agency for Safety and Health at Work estimates a relative increase in highly qualified workers and less demand for less qualified workers performing routine mental tasks or manual tasks, with the disappearance of around one third of currently existing jobs.[8]

In a specific study on robots and jobs carried out by Graetz and Michaels (2015), it was demonstrated that, in industries with high densities of robots, the less qualified workers work fewer hours. Though robotics can affect all the industrial sectors of the economy in different ways, it is also likely that it will affect the jobs within those sectors in different ways.

The World Economic Forum has warned[9] that the digitalisation of industry could lead to the disappearance of 7.1 million jobs and the creation of 2.1 million new jobs. That same report calculates that 65% of today's primary students will have professions that do not even exist at the present time.

For its part, the Observatory for Employment in the Digital Age estimates that 8 out of every 10 people between the ages of 20 and 30 will have digital jobs, which do not exist today.[10]

Those data represent a very serious challenge for today's societies, due not only to the considerable need for transformation and resources that implies but also to the need to adapting education with the aim of minimising the impact of the digitalisation of the economy on both employment and unemployment worldwide.

Because robotisation may mean shrinking demand for workers, especially those who perform routine tasks which can easily and rapidly be mechanised, but also an increase in qualified workers, especially those which require certain creativity or personal or interpersonal skills.

Some authors believe that jobs requiring intermediate qualifications are the ones most at risk of disappearing. Those jobs, which, historically, have

[7] "Opinion on AI: anticipating its impact on work to ensure a fair transition" (own-initiative opinion) (2018/C 440/01).

[8] *A review on the future of work: robotics* (2015) Brussels, p. 3.

[9] World Economic Forum, *The Future of Jobs. Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution* (2016), 13, http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf

[10] Open Future, "What will the new employment be like in Industry 4.0?", 2016, https://www.openfuture.org/es/new/como_sera_el_nuevo_empleo_en_una_industria_40

included accountants, clerks and certain assembly-line workers, are relatively easy to convert into routines. That means that less qualified workers will be forced to engage in activities with lower levels of skills, which, in the medium and long term, will be translated into lower salaries and greater possibilities of losing their jobs.

On the other hand, highly qualified jobs which require the ability to resolve situations, intuition and creativity, and tasks which are carried out "in person" and which require certain flexible social communication skills in the provision of services (attention, treatment etc.) are more difficult to convert into routines, and therefore their survival is assured for longer. As indicated in the ILO document *The Future of Work We Want*, robots and computerisation have, historically, not been able to replicate or automate those tasks.[11]

4.2.2 The New Ways of Working

The new technologies care clearly impacting the ways of working, giving rise to new ways of organising work and production or renewing already-existing ones.

The new digital scenario has generated numerous changes in the internal organisation of work, rendering many jobs obsolete and therefore destined to disappear, while at the same time others are appearing, requiring new skills and aptitudes.

Bonekamp and Sure (2015) explain that there will be a decrease in less qualified jobs, and at the same time we will see an increase in the hiring of workers with higher-level qualifications.

For all of this, it will be necessary to adapt the skills of current workers, while at the same time selecting workers with new profiles, already adapted to the new needs and familiar with the new technological challenges. That work will also require ongoing training in new tools and skills in order to adapt to the new situations.

That ability to adapt will be one of the qualities most demanded by companies in their recruitment processes, and that will inevitably lead to reforms in education, which, as we will see later, will move towards the obtainment of those skills, leaving aside others which will be replaced by technology (Salazar, 2019, p. 58).

[11] National tripartite conference. 28 March 2017. ILO.

The specific characteristics of those new forms of work, technological innovation and competence alter the work methods, leading to less direct contact with the company, which could give rise to changes in work procedures: it will be necessary to redefine who the result is to be reported to, how, how often, the types of communication, the frequency and times of meetings, training and so on.

The new technologies have likewise impacted the forms of employment, and their effects are manifested in the appearance of new work situations and typologies, based on the possibilities offered by telematics and the delocation of jobs, permitting the provision of services without needing to leave home. Those changes imply the need to reassess our usual notions of the place and time of work, and it will even require new types of employment contracts.

That transformation is also being helped by flexible working hours policies or measures such as teleworking, together with new hierarchies, whose structures are more horizontal rather than pyramidal.

All of that leads to new forms of work, such as work on technological platforms or co-working due to a conversion of the organisation of work.

But it is not just the ways in which work is organised. In addition, independent or freelance workers have appeared, who use their own resources to perform tasks independently on different projects, working collaboratively with different companies. They want to manage their own time and do only the work they want, at least initially, because financial needs may determine some of their decisions, as we will see when considering work on digital platforms.

Digitalisation has also had an impact on the automation of some tasks, especially administrative tasks, some of which are now done by the customers themselves (filling in forms, withdrawing money, bookings etc.).

4.2.2.1 Teleworking

Telematics, digitalisation and databases are permitting the delocation of jobs, the simplification of tasks, and even working from home, as in the case of teleworking, in which technology makes it possible to do the work somewhere other than the workplace, using the same tools and incorporating others which serve as substitutes for the most social part of the labour relationship, or the virtualisation of work relationships (replacing physical meetings with the use of chats, video-conferencing systems etc.). That displacement also makes it possible to perform tasks at different times, even in countries in different time zones.

All of that is leading to a reorganisation of the work spaces, giving rise to open spaces without fixed work stations, with people moving around for different projects, and even the appearance of recreational areas for workers to enjoy or for informal meetings, as reported in the study by Thompson (2015).

Taking De las Heras (2020) as a reference, teleworking is a form of organisation of work, which responds to both technological innovation—the new information and communication technologies—and organisational flexibility in a changing environment.

Teleworking is not suitable for all kinds of work; it is only possible for those tasks, which can be done through technological means.

The implementation of teleworking in the organisation not only means working from home, it also makes it necessary to redefine some of the elements of that work, such as:

- Time: determining the working hours and the length of the workday, holidays, days off, leave and times of reversibility.
- Space: establishment of the workplace where the task will be done, which must meet certain minimum conditions in order to ensure the effective prevention of occupational hazards and adequate health and safety conditions.
- Remuneration: determining which costs and expenses will be borne by the company and which by the workers, while at the same time it will be necessary to review the complements deriving from work at home.
- Work conditions: health and safety in the workplace.
- Work tools: ownership and maintenance of the work equipment, insurance and the confidentiality and security of data.
- Work methods and training: company control.

4.2.2.2 Work on Digital Platforms

The technological change we are experiencing presages changes in the ways we understand work (Mercader Ugina, 2017, p. 14). Thus, among others, new types of companies emerge, characterised by the use of digital platforms which put business owners, workers and customers in contact with each other in transactions of goods and services.

The use of electronic devices has become widespread, becoming almost basic necessities for citizens and even for some companies, especially those whose activity is based on e-commerce. It is in that environment that digital platforms emerge as meeting places for suppliers and users, where they share, exchange or provide goods and services.

The example of digital platforms perfectly fulfils that new reality: their algorithms control and monitor the workers, and assess their performance, while offering workers the opportunity to work from any place, at any time and in any job which is suited to them. In the words of Tirole (2017, p. 443), on digital platforms "everyone is a winner".

In particular, digital work platforms are opening up opportunities, which previously did not exist, especially for women, young people, people with disabilities and marginalised groups throughout the world.

With digital platforms, it becomes possible to divide the work into micro-tasks (short individual tasks), which means contracts for the exact time and the specific duration of the service, and for minimum remuneration.

The majority of the time, workers on platforms are free to choose their own timetables, their availability, the place they work from and so on. It could be said that they are their own bosses.

However, there are not only advantages to that "on-demand" digital work. It can also generate considerable health risks deriving from the harsh working conditions, low salaries and job insecurity, together with the loss of privacy due to the traceability of each one of the operations by those in charge, and even by customers. Likewise, those customers may be able to score each activity in order to generate confidence among other customers, but that may also cause a loss of the worker's income and reputation.

Likewise, it permits greater surveillance by the company, which can become invasive, due to the many opportunities afforded by video-surveillance, geolocation, biometric controls, the monitoring of emails and Internet browsing, or the way in which social networks enable companies to monitor and control their workers.

4.2.3 The New Professions in Industry 4.0

The application of the new technologies to surveillance and control of the work carried out, the use of email, access to the Internet, the automated processing of workers' data and information, the acceptance of electronic signatures and the use of the intranet in collective relations have all had a great impact on the world of work, with the appearance of new forms of work (Blasco Pellicer, 2009).

Though it is a common belief that technology will take over many jobs, which will lead to massive unemployment (Berg et al., 2016, p. 10), so far it seems unlikely that that scenario will occur (Arntz et al., 2016; Frey & Osborne, 2017; Manyika et al., 2017), in that many activities still cannot be carried out by machines and machines are not capable of doing different activities in an integrated manner. Up to now, in general, it has only been possible to replace individual activities within jobs.

There are two reasons for that. First of all, many middle-rank jobs require a combination of non-routine tasks and capabilities, including emotional skills, problem-solving and flexibility, which cannot, as yet, be done by machines. Secondly, the rise of new technologies has created new intermediate-rank jobs, such as health-care assistants, and it has stimulated the demand for others, such as managers of food establishments. Consequently, machines are increasingly capable of providing services which require few qualifications and highly qualified cognitive work (Holzer, 2015).

The technical positions in which there is expected to be growing demand in all sectors include data analysts, scientists, software and applications developers, and specialists in e-commerce and social networks, and at the same time there will be a demand for profiles capable of innovating and programming, those with leadership qualities, more artisanal occupations or those who produce unique pieces which technology is not capable of matching.

4.2.4 Digitalisation and Worker Well-being Versus New Psychosocial Risks in Work

Automation and the incorporation of robots into the workplace could minimise the occupational hazards to which workers are exposed, due to less exposure to dangerous substances and accidents, avoiding repetitive tasks or those, which are harmful to our health.

Thus, exoskeletons can enable older workers and those with reduced mobility or physical disabilities to do work which requires physical effort, while drones avoid work at heights, travel across large areas or entry into difficult or dangerous places (Rimbau-Gilabert, 2019, p. 9).

In short, robots are key in relieving physical fatigue and the hardships of many jobs in both industry and the services sector, and in improving the well-being of workers.

At the same time, digitalisation and robotisation may generate negative consequences for workers by reducing the variety of tasks to be carried out, focussing on more complex activities or those which require greater skill, requiring greater effort by workers (Stacey et al., 2018), which could have a negative impact on their mental health or the appearance of muscular-skeletal problems caused by reduced physical activity and more static postures, while at the same time diseases associated with the use of technology appear, such as technostress, which is defined by the Framework Agreement on Stress at Work (2004) as "a state which is accompanied by physical, psychological or social complaints and which results from individuals feeling unable to meet the requirements or expectations placed on them".

Likewise, the interaction between robots and workers must be studied in order to avoid knocks or accidents, and also to ensure that there is no violation of workers' rights or even other aspects of people's daily lives when they come into contact with the use of technology (Aguilera Durán, 2019), or, as the European Economic and Social Committee warned, in the official opinion on AI at the end of 2018, it is necessary to "draw a line in the sand for interaction between workers and intelligent machines so that humans never become the underlings of machines".[12]

4.3 The Impact of AI on Education

AI will profoundly transform education, declared Audrey Azoulay, Director General of the UNESCO at the UNESCO Mobile Learning Week.[13]

AI brings with it the need to transform education systems, from their methods to the way in which knowledge is accessed, and the training of teachers is fundamental in that evolution.

It means updating regulated academic education (primary, secondary, university and professional training) to meet the needs of the labour market and also those of a digitalised society.

According to the study entitled "Artificial intelligence and life in 2030", carried out by Stanford University, California, USA, the use of AI will be

[12] Opinion of the European Economic and Social Committee on "Artificial intelligence: anticipating its impact on work to ensure a fair transition" (2018/C 440/01).

[13] World Federation of Engineering Organizations. (2019). *Report on Mobile Learning Week 2019*. Retrieved from: https://www.wfeo.org/report-on-mobile-learning-week-2019/, carried out in March 2019 in Paris France.

habitual in classrooms in the near future, together with other technologies such as virtual reality, educational robotics or smart tutoring systems. That use will, as indicated by Salazar (2019, p. 74), permit the personalisation of teaching, because, through different tools and the gathering and analysis of data, we will be able to know each student's specific needs, thereby facilitating individualised education, which maximises the skills and minimises the weaknesses.

Moreno Padilla (2019) shares that opinion, considering that the aim of AI in education must be the creation of programmes which permit adaptive, personalised learning environments, as in the case of chatbots or tools which make it possible to simulate conversations with people by voice or via the keyboard. That tool is starting to be used in automatic assessments, which permit both synchronous correction and the generation of different evaluation models through the answers given by the student.

But the Education System must not only incorporate technological subjects, it must also include subjects which make it possible to develop skills which cannot be automated and which will be those sought after by companies in their recruitment processes, which may include the following (Magro et al., 2014):

- **Digital knowledge**: making your way, personally and professionally, in the digital economy.
- **Management of information**: seeking, obtaining, evaluating, organising and sharing information in digital contexts.
- **Digital communication**: communicating, interacting and collaborating efficiently with tools and in digital environments.
- **Online working**: working, collaborating and cooperating in digital environments.
- **Continuous learning**: managing learning autonomously, knowing and using digital resources and maintaining and participating in learning communities.
- **Strategic vision**: understanding the digital phenomenon and incorporating it into the strategic orientation of the organisation's projects.
- **Online leadership**: directing and coordinating distributed work teams online and in digital environments.
- **Customer focus**: listening, understanding, knowing how to interact with and meet the needs of the new customers in digital contexts.

In addition to those new digital skills, the work of the future will require traditional professional skills such as creativity, complex problem-solving, critical thinking, people management, coordination, decision-making and a focus on customers and service.[14]

That situation will require different qualities and qualifications of human resources so they can adapt to the new ways of organising the work and the new ways of doing things, requiring an effort on the part of not only companies but also governments in order to make adjustments to education and professional training policies.

That evaluation of skills more than academic training will make it necessary to adapt the recruitment processes, incorporating tools which make it possible to detect the personal skills, seeking to improve the performance of employees, or a strategy to retain talent. A particularly highly valued attitude will be the ability to adapt to changing circumstances and new relations.

The business agility cites the following examples of the application of AI in Human Resources[15]:

- **Improving performance**. AI is very useful in any company department, which works with data and where decisions are based on objective analyses. Thanks to its application, it is possible to improve the organisation of work in accordance with the performance of the workforce.
- **Analysing absenteeism**. Thanks to analysis of the information related to worker absences, it is possible to predict when absenteeism from work is more frequent, in order to both mitigate its effects and develop and implement an improvement plan.
- **Company selection and recruitment processes**. AI can be applied to selection tests in order to identify those candidates who meet the established requirements and who are likely to perform best within the organisation. It is based on pre-established patterns and the behaviour of the candidates during the tests, reducing the relative importance of the person's CV.
- **Retaining talent.** Algorithms also play a part in determining the characteristics of high-performance teams, which employees have the

[14] Randstad, Digitalisation: does it create or destroy employment?, Randstad Report 2016.

[15] Examples of AI application in Human Resources available from 5 ejemplos de aplicación de Inteligencia Artificial en RRHH | aggity.

greatest leadership capabilities, or which ones are most likely to leave the organisation.

- **Designing training plans.** Equipping employees with new skills is essential for both their personal development and that of the company. Knowing how frequently training initiatives are carried out and what their results are helps in planning in accordance with the needs of the company, and also to ensure that it is an incentive.

But not only will those tools help to detect qualities in workers, they will also be used to bring together the different sources of information which the recruiter currently has—not only is the candidate's CV evaluated but also their professional and general social networks can be consulted.

To carry out any of those actions, the company must have concrete, measurable data, and it must take into account only those data, which are truly relevant, and in those tasks human intelligence is also important, at least for the time being.

Likewise, big data is used in HR departments for staff management, for tasks such as keeping records of work times, especially in large companies and those in which there are shifts and flexible working hours, or systems to assess and supervise performance or the establishment of work schedules in accordance with the worker's availability.

But it is not only necessary to take into account the skills of workers at the start of the labour relationship, it will also be necessary to maintain those skills, incorporating flexible, agile training which enables workers to adapt to the changes, especially in those jobs which could be replaced by technology.

References

Aguilera Durán, J. (2019). *Right to Work, Labor Automation and Rights of Affectation by Technology's Use*. Revista Latinoamericana de Derecho Social, no. 29, Universidad Nacional Autónoma de México, Mexico. http://www.redalyc.org/articulo.oa?id=429662331001; https://doi.org/10.22201/iij.24487899e.2019.29.13898.

Arntz, M., Gregory, T., & Zierahn, U. (2016). *The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis*. OECD Social, Employment and Migration Working Papers, No. 189. Paris.

Berg, A., Buffie, E. F., & Zanna, L. F. (2016). Robots, Growth, and Inequality. *Finance & Development, 53*(3), 10–13.

Bersin, J., Flynn, J., Mazor, A., & Melián, V. (2017). *Rewriting the Rules for the Digital Age. Global Human Capital Trends.* Deloitte University Press, 1–15.

Blasco Pellicer, C. (2009). Incidencia de las nuevas tecnologías de la información y la comunicación (TICS) en las reestructuraciones de las empresas, *Revista Aranzadi Social*, 2(15) (Study).

Bonekamp, L., & Sure, M. (2015). Consequences of Industry 4.0 on Human Labour and Work Organisation. *Journal of Business and Media Psychology, 6*(1), 33–40.

Cedrola Spremolla, G. (2019). Competencias, Organización del Trabajo y Formación Profesional en el Trabajo del Futuro: Algunas reflexiones para posibilitar cambios imprescindibles. *Revista Internacional y Comparada de Relaciones Laborales y Derecho del Empleo, 7*(1). ADAPT University Press, 6–51.

European framework agreement on harassment and violence at work /* COM/2007/0686 final */EUR-Lex - 52007DC0686 - EN (europa.eu)

Framework Agreement on Stress at Work. (2004). European Framework Agreement on Harassment and Violence at Work. COM/2007/0686 final. EUR-Lex – 52007DC0686 – EN (europa.eu).

Frey, C. B., & Osborne, M. A. (2017). The Future of Employment: How Susceptible are Jobs to Computerisation? *Technological Forecasting and Social Change, 114*, 254–280.

Graetz, G., & Michaels, G. (2015). *Robots at Work*. Centre for Economic Performance Discussion Paper No. 1335, London School of Economics.

Heras García, A. de las (2020). Implantación del teletrabajo ante una situación de emergencia. *CEFGESTION*, 260 CEF, 110–120.

Hislop, D., Coombs, C., Taneva, S., & Barnard, S. (2017). *Impact of Artificial Intelligence, Robotics and Automation Technologies on Work*. Chartered Institute of Personnel and Development.

Holzer, H. (2015). *Job Market Polarization and US Worker Skills: A Tale of Two Middles. Economic Studies.* The Brookings Institution.

Magro, C., Salvatella, J., Álvarez, M., Herrero, O., Paredes, A., & Vélez, G. (2014). Cultura Digital y Transformación de las Organizaciones. 8 competencias digitales para el éxito profesional, RocaSalvatella, at http://www.rocasalvatella.com/es/8-competencias-digitalespara-el-exito-profesional. See also Rapport Mettling, Transformation numérique et vie au travail, and also Brynjolfsson, E., McAfee, A., *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* New York: Norton & Company. Retrieved from https://tanguduavinash.files.wordpress.com/2014/02/thesecond-machine-age-erik-brynjolfsson2.pdf

Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). *Harnessing Automation for a Future that Works.* McKinsey Global Institute.

Mercader Ugina, J. (2017). La robotización y el futuro del trabajo. *Trabajo y Derecho, 27*, 14.
Moreno Padilla, R. D. (2019). La llegada de la IA a la Educación (The Arrival of Artificial Intelligence to Education). *RITI Journal, 7*(14), 262. https://doi.org/10.36825/RITI.07.14.022
Pérez-González, D., Solana-González, P., & Trigueros-Preciado, S. (2018). Economía del dato y transformación digital en PYMEs industriales: retos y oportunidades. *Economía Industrial, 409*, 37–46.
Rimbau-Gilabert, E. (2019). *Digitalización y bienestar de los trabajadores.* IUSLabor 2/2019, p. 9.
Salazar, I. (2019). *La revolución de los robots: cómo la IA y la robótica afectan a nuestro futuro*. Gijón, Ediciones Trea. 1–169.
Schlogl, L., & Sumner, A. (2018). *The Rise of the Robot Reserve Army: Automation and the Future of Economic Development, Work, and Wages in Developing Countries.* Working paper 478. Center for Global Development. Washington.
Stacey, N., Ellwood, P., Bradbrook, S., Reynolds, J., Williams, H., & Lye, D. (2018). *Foresight on New and Emerging Occupational Safety and Health Risks Associated with Digitalisation by 2025* [Online]. https://osha.europa.eu/en/tools-and-publications/publications/foresight-new-and-emerging-occupational-safety-and-health-risks/view
Thompson, P. (2015). Nuevas formas de trabajar en la empresa del futuro. In *Reinventar la empresa en la era digital.* Open Mind BBVA.
Tirole, J. (2017). *La economía del bien común*. Taurus.
Valenduc, G., & Vendramin, P. (2016). *Le travail dans l'économie digitale: continuités et ruptures.* Working Paper ETUI, no. 3.

CHAPTER 5

AI Integration in the Digital Transformation Strategy

Ana Landeta Echeberria

Abstract The speed of Artificial Intelligence (AI) adoption today varies across companies and industries, as it requires complementary research in data, skills, digitization of workflows and the capacity to adapt organizational processes. This makes AI and knowledge management for managerial decision-making highly relevant, in addition to awareness of how to develop a Digital Transformation Strategy Plan within a Digital Business Strategic Framework. All of this is necessary and may be adapted to any company and sector.

However, AI has intensified the need for this approach if we are to create new opportunities for digital transformation in companies as well as new challenges for managers of digital transformation processes.

This chapter presents three frameworks (Digital Strategic Framework, AI readiness and AI integration in Digital Transformation Strategy Model) that help to address new businesses challenges and a series of strategic-operational suggestions that permit an initial and future approach

A. Landeta Echeberria (✉)
Faculty of Economics and Business, Universidad a Distancia de Madrid (UDIMA), Madrid, Spain

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_5

associated with AI integration, as new technological elements of the Digital Transformation Strategy.

Keywords Knowledge management • AI readiness • Digital transformation • Business strategy • Strategy framework

5.1 AI and Knowledge Management and Managerial Decision-Making

At current the business scenario, AI ethics and its impact on knowledge management is a fundamental question that is studied by A. J. Rhem (2021) deeply as follows.

Many companies and sectors lag in AI adoption. Developing an AI strategy with clearly defined benefits, finding talent with the appropriate skill sets, overcoming functional silos that constrain end-to-end deployment and lacking ownership and commitment to AI on the part of leaders are among the barriers to adoption most often cited by executives.

On the strategy side, companies will need to develop an enterprise-wide view of compelling AI opportunities, potentially transforming parts of their current business processes. Organizations will need robust data capture and governance processes as well as modern digital capabilities, and be able to build or access the requisite infrastructure. Even more challenging will be overcoming the "last mile" problem of making sure that the superior insights provided by AI are inculcated into the behaviour of the people and processes of an enterprise.

Strategies and means for selecting and implementing digital technologies that realize firms' goals in digital transformation (DT) have been extensively investigated. The recent surge in AI technologies has amplified the need for such investigation, as they are being increasingly used in diverse organizational practices, creating not only new opportunities for digital transformation but also new challenges for managers of digital transformation processes. In this chapter I present a framework intended to assist efforts to address one of the first of these challenges: assessment of organizational AI readiness, that is, an organization's ability to deploy AI technologies to enable digital transformation, in four key dimensions: technologies, activities, boundaries and goals. I show that it can facilitate analysis of an organization's current socio-technical AI status, and future prospects for its fuller value-adding socio-technical deployment. The AI readiness framework invites fuller theorizing of the roles that AI can, and will, play in digital transformation.

AI plays an important part to delivering knowledge in a digitized organization by elevating how the delivery of knowledge occurs to the people who need it. At the same time, AI will facilitate the mapping of knowledge (tacit and explicit) to enable everyone seeking knowledge within the organization to be provided its necessary context. As knowledge within the organization grows, fast response and personalized access to organizational knowledge assets have become necessary to execute and deliver results. This is particularly important for content-related products and services, such as consulting, marketing/advertising and communications within an organization. The digital nature of these products allows for more customized delivery within the AI Framework. To provide personalized services a complete understanding of user profiles and accurate associations to people are essential.

5.2 A Digital Strategic Framework[1] (Elements and Phases)

5.2.1 *The Relevance of Business Models and the Existence of a Digital Strategy*

In this new age of knowledge and digital economy, business models are assuming increasing importance as elements of strategic innovation and main driving forces to firm's competitiveness. Those circumstances are creating new demands to understand how firms design and systematize their operations, generate value for costumers and draw the costs structure, revenues and profits to be distributed, all elements of a business model (Chesbrough & Rosenbloom, 2002; Teece, 2010; Zott & Amit, 2013). But also as firms address the characterization and measurement of innovation practices (Adams et al., 2006) and endorses the complexity of business modes in relation to strategy and the management of innovation emerging from IT (Powell, 1992; Tidd, 2001, 2006; Bergeron et al., 2004).

Managing digital innovation has become an important issue not only for contemporary firms, by nature more open for innovation, but for all firms, namely those living a digital transformation process (Holmstrom & Nyle, 2015).

[1] The full version of the work "A Digital Framework for Industry 4.0 Managing Strategy" (Echeberria, L., 2020)" is published at https://link.springer.com/book/10.1007%2F978-3-030-60049-5

Bharadwaj et al. (2013), underlining the rapid growth of a digital world during the last decade, assumed the existence of a digital business strategy which is inherently simultaneously strategic and technological due to the ubiquitous presence of digital in all dimensions of decision, overlapping business strategy with IT strategy.

They have defined digital business strategy has an "organizational strategy formulated and executed by leveraging digital resources to create differential value" (Bharadwaj et al., 2013, p. 472).

Harmancioglu et al. (2009) introduced the concept of resource fit, which refers to strategic fit resources for a new initiative. Based on this perspective, digital transformation begins with the evaluation of resources so as to generate a firm's competitive advantage and the identification of sources of synergy or fit. Apparently, resource fit provides a broadened view that considers both resource-based theory and strategic fit view to extend our understanding of digital transformation. However, to date, little research has exploited the concept of resource fit in digital transformation; thus, the critical factors for making a successful digital transformation are still largely unexplored.

5.2.2 Construct of Strategic Framework

This strategic framework is composed of a research model, elements, phases of development, actions and activities, which will be described below.

5.2.2.1 Research Model

Due to the nature of this research, a methodological approach based on a combination of synthetic and analytical methods was chosen with a view to develop a Digital Transformation Research Strategy Model for companies within the Industry 4.0 framework.

Said combination will facilitate greater understanding of the phenomenon and its subsequent application. This focus adjusts to the research as it is intended to investigate how to prepare a Digital Transformation Strategy Framework within the context of Industry 4.0. The confirmation of a digital strategy in the context of companies with the aim of digitizing existing business models and/or implementing improvements to the existing one is a relatively recent phenomenon that continues to develop within different industrial sectors.

In this sense, the study primarily uses this methodological combination in which, data, trends and policy initiatives collected will be used as a source from which conclusions can be drawn that may drive the development of a specific strategic focus.

The synthetic method is a reasoning process that tends towards rebuilding the whole based on the elements distinguished by the analysis. In short, it consists of a methodical and brief explosion. In other words, we must say that synthesis is a mental process aimed at obtaining a thorough understanding of the essence of know-how in all its parts and particularities. Synthesis means reconstructing and re-integrating the parts of the whole. However, this operation implies overcoming the analytical procedure, as not only does it represent the mechanical reconstruction of the whole, as this will not allow for the advancement of knowledge but that it implies the understanding of the essence of same, to ascertain the basic aspects and relations from the perspective of the totality. There is no synthesis without analysis, as Engel (1999) said, as analysis provides the raw material with which synthesis can be performed.

The analytic-synthetic model is a route to knowledge as it involves analysis (from the Greek *análysi*, meaning "decomposition"), that is, the separation of a tone in its parts and constitutive elements. It is supported, therefore, on the belief that in order to know a phenomenon it is necessary to break it down into its parts. In contrast with that manifested, the synthetic method involves synthesis (from the Greek *synthesis*, meaning "union"), that is to say, the union of elements to form a whole.

Thus, the resolution method with heuristic content since in Descartes's own formulation (Descartes, 1998), in his "Regulae ad directionem ingenii and in his Geometric" (Descartes, 1983), the method can be formally presented, broken down into a series of steps, which the resolver literally goes through: (1) a heuristic reading of the problem statement that reduces it to a list of quantities and relationships between quantities; (2) choice of a quantity to be represented with a letter (or of a few quantities that are going to be represented with different letters); (3) representation of other quantities by algebraic expressions that describe the relationship (arithmetic) that these quantities have with others that have been previously represented by an algebraic letter or expression; (4) establishment of an equation (or so many different letters that have been decided to introduce in the second phase), matching two expressions, of those described in the third step, that represent the same amount; (5) transformation of the equation in a canonical way; (6) application of the solution formula or algorithm to the equation in a canonical way; and (7) interpretation of the result of the equation in terms of the problem.

The analytical judgement involves the decomposition of the phenomenon into its constitutive parts. It is a mental process that divides the complete representation of a phenomenon into its parts.

The synthetic judgement, on the contrary, consists of systematically joining the heterogeneous elements of a phenomenon with the purpose of rediscovering the individuality of everything observed. Synthesis means the act of unifying the disparate parts of a phenomenon. However, synthesis is not the sum of the partial contents of a reality: synthesis adds to the parts of the phenomenon something that can only be acquired on the whole, in its singularity.

Consequently, and taking the methodological focus mentioned previously and the nature of this work into account, the construction of the analytic-synthetic model combines the assembly of a tool to measure the level of maturity with an evolution in accordance with the diagnosis from a strategic perspective driven towards the preparation of a development plan for the digital transformation of companies.

5.2.2.2 Methodology

The design of the analytical and systematic research methods of this work is illustrated in Fig. 5.1.

Due to the nature of this research, a methodological focus based on new, exploratory research was chosen. This research also intends to provide a general and approximate view in order to develop an Internal

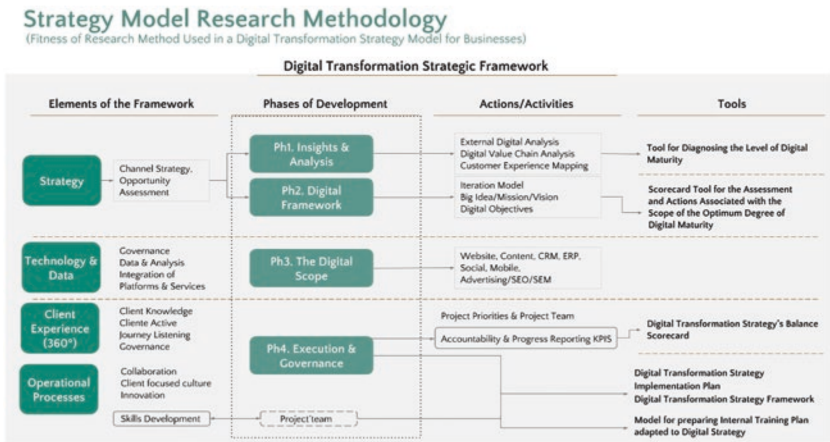

Fig. 5.1 Research model and methodology. (Source: Author's own)

Training Plan model adapted to the digital strategy of the company within the framework of Industry 4.0.

As a barely explored, emerging area and, furthermore, one where the nature of the subject of study does not allow for the formulation of a precise hypothesis, Industry 4.0, understood as the fourth industrial revolution, is a new phenomenon that is so new that it does not allow a systematic description at present. So, resources for researchers are still insufficient to undertake more in-depth work.

5.2.2.3 Fitness of Research Method Used in a DT Strategy Model for Businesses

The objective of the DT Strategy is to create the capacity necessary for taking maximum advantage of the possibilities and opportunities offered by new technologies and their impact in a more rapid, better and more innovative way in the future. A DT journey needs a phase-based focus with a clear roadmap that involves a variety of interested parties, beyond the silos and limitations, internal and external. This DT Strategy Framework takes into account the fact the final objectives will continue to advance given that digital technology is a continuous journey, just like digital change and innovation are.

Therefore, the DT Strategy is the process of identifying, articulating and executing digital opportunities that extend the competitive advantage of the organization.

So, if the DT Strategy is a process, it cannot be forgotten that a process represents progress, from an initial starting point, to the approaching and surpassing landmarks, both identified and unknown as well as with measurement. Therefore, the proposed solution goes through the application of the metric presented below.

In order to build up a *research method used in a DT strategy model for businesses*, as a starting point the doctoral thesis “A Digital Transformation Strategy Model for companies within the Industry 4.0 framework” by the same author as the current study was taken.

Therefore, the synthetic model research contains a review of the scientific literature mentioned above, mainly topics related to sociocultural and political transformations of the Information Society, the Industrial Internet and the potentially economically disruptive technologies, skills and the jobs of the future and the New Economy and new business models.

Secondly, the parts components of the strategic framework were defined: definition of DT, measuring of digital maturity model and generation of digital maturity evaluation level tool, associating them with processes inherent in the construction of a plan of digital transformation of a part.

Third, following the guidelines of the Analytical Method (systematic union of heterogeneous elements; specifically, types of diagnoses associated with the digital maturity evaluation level and assessment scorecard) we have resulted in the main elements of the DT Strategic Framework; Strategy, Technology & Data, Client Experience (360°) and Operational Processes.

The elements, phases of development, actions and tools related to the DT Strategy Framework will include those presented graphically in Fig. 5.2.

5.2.3 *Phases of Development*

The Phases and Actions/Activities interrelation of DT Strategy Framework will result in the creation of a Strategic DT Operational Framework complementary to DT Strategic Framework.

The aforementioned Phases are detailed below.

- PHASE 1. INSIGHTS AND ANALYSIS: understanding the needs and priorities of the people who are at the core of your digital revolution including your customers, stakeholders, employees and

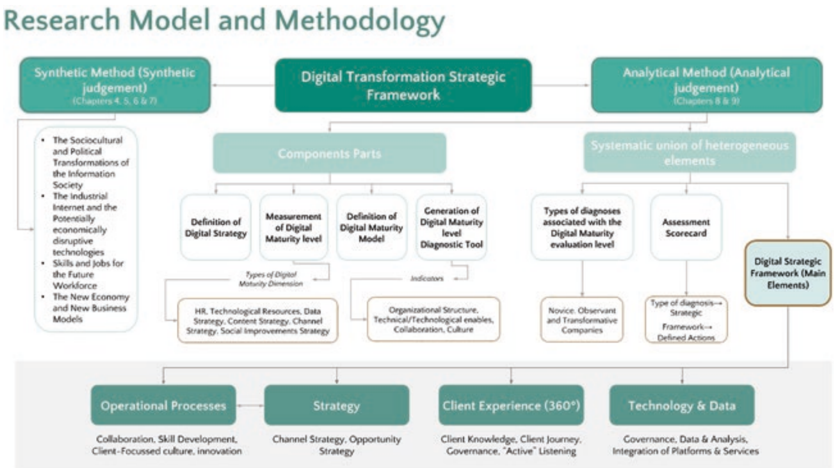

Fig. 5.2 Strategy model research methodology; the DT Strategy Framework (interrelation of elements, actions, activities, phases and tools). (Source: Author's own)

executives. Analyse internal performance and sales data to understand where the biggest value is External Digital Analysis, Customer Experience Mapping and Digital Value Chain Analysis.

- PHASE 2. DIGITAL FRAMEWORK: creating a framework that allows the company to addresses digital goals and objectives; The Iteration Model, Big Idea/Mission/Vision, Digital Objectives and Unique Value Proposition (UVP).
- PHASE 3. THE DIGITAL SCOPE: addressing the company's approach to key areas of Digital and outlining the purpose, objectives and key initiatives and challenges of each channel; Website(s), Online Content, Digital Advertising/ Search Engine Marketing (SEM)/ Search Engine Optimization (SEO), CRM, Social, Mobile and Enterprise Resource Planning (ERP).
- PHASE 4. EXECUTION & GOVERNANCE: prioritizing the plan, taking into account which needs are the most urgent and important as well as current resources, timelines and budgets. Once digital business models become more agile we're forced to create plans that are iterative. Unlike traditional business planning, looking 3–5 years ahead is rarely realistic and accurate budgeting can be nearly impossible; Project priorities, Project team(s) and Accountability and Progress Reporting/ Key Performance Indicator (KPIs).

And, the interrelations between Phases, Tools, Patterns, Actions, Approaches and Measures derive in solutions understood as tools of different utility during the development of the phases mentioned above, as can be seen in Fig. 5.3.

5.2.4 Tools

Therefore, in parallel to the DT Strategic Framework, this research work presents as one of the contributions, the construction of a Strategic DT operational framework, necessary and adaptable to any type of company and sector of activity (Fig. 5.4).

Moreover, the DT Strategic Framework is a compendium of tools build ad hoc that will allow the development of the actions inherent to the conformation of the Business DT Plan. Furthermore, carried out together with the DT Strategy Framework, both frameworks will ensure a Business DT Plan design and execution with grants from an operational and strategic perspective, respectively.

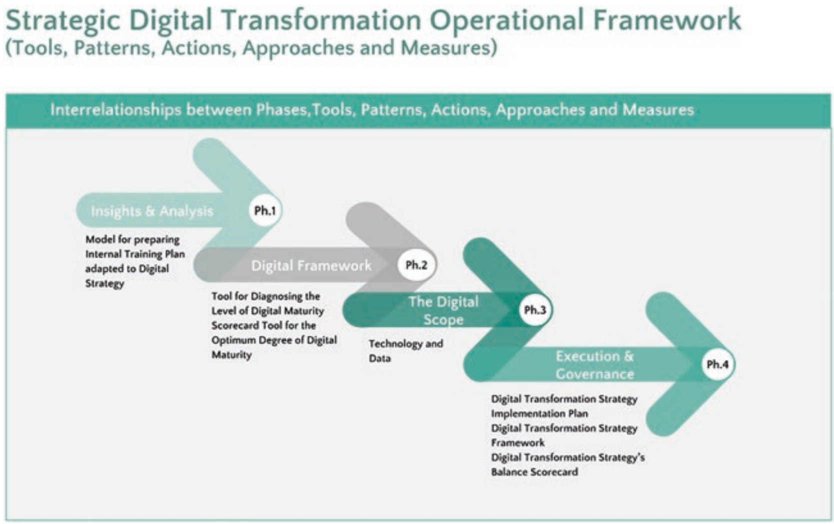

Fig. 5.3 Digital Transformation Strategy Framework (phases and actions/activities interrelation). (Source: Author's own)

Consequently, the strategic framework suggested includes the patterns, actions, approaches and several measures that are detailed below and fully explained in "A Digital Framework for Industry 4.0 Managing Strategy" (Echeberria, 2020) (useful documentation for companies not interested in knowing the results of this work from a purely academic perspective).

- Model for preparing Internal Training Plan adapted to DT Strategy.
- Tool for diagnosing the level of digital maturity and a scorecard tool for the assessment and actions associated with the scope of the optimum degree of digital maturity.
- Digital Transformation Strategic Framework (elements and phases).
- Digital Transformation Strategy Plan (phases and actions).
- Digital Transformation Strategy's Balance Scorecard.

5.3 Digital Transformation Strategy Framework

The objective of the Digital Transformation Strategy is to create the capacities necessary for taking maximum advantage of the possibilities and opportunities offered by new technologies and their impact in a more

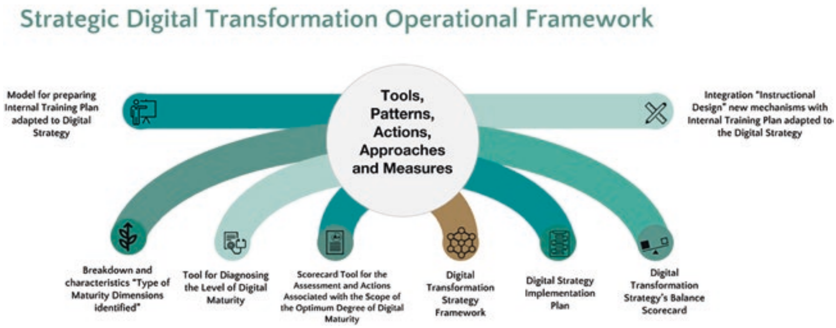

Fig. 5.4 Strategic Digital Transformation Operational Framework. (Source: Author's own)

rapid, better and more innovative way in the future. A digital transformation journey needs a phase-based focus with a clear roadmap that involves a variety of interested parties, beyond the silos and limitations, internal and external. This Digital Transformation Strategy Framework takes into account the fact the final objectives will continue to advance given that digital technology is a continuous journey, just like digital change and innovation are.

In order to develop the Digital Transformation Strategy Framework, it is considered appropriate to have as a reference to Harrison's (2015) "The Digital Strategy Guide". In Harrison's Master Plan, the Digital Strategy consists of four main parts that which we will summarize and concretize according to the research approach proposed in the present work.

5.3.1 *Phase 1: Insights and Analysis*

Understanding the needs and priorities of the people who are at the core of your digital revolution including your customers, stakeholders, employees and executives. Analyse internal performance and sales data to understand where the biggest value is.

- External Digital Analysis
- Customer Experience Mapping
- Digital Value Chain Analysis

5.3.2 *Phase 2: Digital Framework*

Creating a framework that allows the company to addresses digital goals and objectives.

- The Iteration Model
- Big Idea/Mission/Vision
- Digital Objectives
- UVP

5.3.3 *Phase 3: The Digital Scope*

Addressing the company's approach to key areas of digital and outlining the purpose, objectives and key initiatives and challenges of each channel.

- Website(s)
- Online Content
- Digital Advertising/SEM/SEO
- CRM
- Social
- Mobile
- ERP

5.3.4 *Phase 4: Execution and Governance*

Prioritizing the plan, taking into account, which needs are the most urgent and important as well as current resources, timelines and budgets. Once digital business models become more agile we're forced to create plans that are iterative. Unlike traditional business planning, looking 3–5 years ahead is rarely realistic and accurate budgeting can be nearly impossible.

- Project priorities
- Project team(s)
- Accountability and Progress Reporting/KPIs

In substance, and in a broader sense, the Digital Strategy' four main parts are characterized by:

5.3.4.1 Phase 1: Insights and Analysis

Research is time-consuming and expensive, but it's nothing compared to the cost of a project that failed due to lack of diligence.

External Analysis (the Digital Way)

Since the Digital Strategy is about examining the business model as a whole, it is recommended starting with an external analysis like Porter's (2001) Five Forces: threat of new entrants, determinants of supplier power, rivalry among existing firms, determinant of buyer power and threat of substitutes. The exercise allows you to thoroughly consider and evaluate the digital threats and opportunities that could come from outside the business.

To keep the exercise focused, only analyse the Five Forces as they relate to the digital business model.

A breakdown of how it works is as follows:

- **Rivalry Among Existing Firms**. To identify the threats and opportunities of direct competitors.
- **Threat of New Entrants**. To think about the start-ups and companies that could become threats.
- **Threat of Substitutes**. To be on the lookout for new business models here.
- **Determinants of Supplier Power**. To consider what your suppliers are doing digitally but it could also mean that the power that your digital suppliers have over the company.
- **Determinants of Buyer Power**. To consider the power your digital buyer has and how the company can mitigate the risk involved in the customer relationship.

This factor is especially relevant to small to mid-size business units whose buyers are larger than they are.

Customer Experience Mapping

Designing your customers' experience starts with understanding and empathizing with their pain. Hopefully, you already have customer/user personas that represent your key targets. These personas shouldn't just be in a folder on your desktop—I recommend having them printed and on your wall so they're always top-of-mind.

Main aim: to use a journey mapping to outline the high-level experience of the customer life cycle, then break it down into smaller pieces and layer in other components like sales activity, brand touch-points and user experience for the company website or app.

Journey mapping is an essential part of your Digital Strategy because it forces you to empathize with all of your users and stakeholders both internal and external. It also reduces confusion by making the process crystal clear.

Digital Value Chain Analysis

To do a traditional internal analysis of the value chain following Michael Porter's theory. Like the external analysis, we'll put a digital twist on it.

To keep in mind that this *does* apply to service businesses just as much as products. Language like "inbound logistics" suggests that we're only talking about a product business—but truthfully the process is just the same.

Starting with primary activities, take each business unit and break down its strengths and weaknesses.

- **Inbound Logistics**. To analyse the digital process of procuring goods from suppliers in the value chain.
  To embrace dynamic ecosystems, in order to connect the company and suppliers together to rapidly and automatically respond to changing conditions.
  This digital ecosystem creates a dynamic supply chain, all the way from raw materials' providers to end consumers, according to VanBoskirk et al. (2017).
  To evaluate the digital aspects on the front end of your value chain.
- **Operations**. To have an operations team happy with the software they are dealing with.
- **Outbound Logistics**. To evaluate the technology you use in your distribution processes. Extending products and services with digital technologies makes it possible to take greater mindshare within a customer's ecosystem of value, according to VanBoskirk et al. (2017).
- **Marketing & Sales**. To focus on customer data as a big consideration. Taking into account that Marketing and Sales often need to leverage many online platforms for CRM, Content Management, E-Commerce, Business Intelligence, Lead Nurturing and much

more. The Big Data movement is hitting these customer-facing units hard and the way you leverage company's customer data should be regularly evaluated.

- **Service.** To focus on what happens after the purchase: the customer's service relationship with the brand long term and to provide them with what they need to be acquired digitally.

5.3.4.2 Phase 2: Digital Framework

Main aim: to create a framework that allows the company to addresses digital goals and objectives. To cast an overall vision for what digital will be doing for the company. This includes one big idea and your digital objectives.

(a) **Create Your Iteration Model**. To consider the digital transformation as an evolutionary process—not a traditional business plan where we plan then execute over long periods of time. In fact, most digital work should happen in six-week sprints where company should be focusing all the energy on one important short-term goal.

(b) **Distilled to One Big Idea**. Since your Digital Strategy is all about people, you *need* people to believe in your ideas and adopt them as their own. To distil the entire plan to one big idea that everyone can remember and believe in.

Examples:
We are

- Turning visitors into customers
- Going from "great" to "irreplaceable"
- Becoming our customers' favourite company

To circle back to this exercise at the end of your process.

(c) **Digital Base: Your Goals and Objectives**. To define what the digital program to do for the company. To describe both the goal and the objectives that will be used to complete it.

Here are some definitions for the purpose of this exercise:
Digital Goal: This crystalizes the thing that stakeholders want most.
Example:

Goal 1: To increase mobility for national sales force and account managers, giving them visibility into the sales pipeline, account activity and company performance.

Objectives:

To seek and implement a CRM platform that is accompanied by a mobile app to capture and view lead data on the road.

To find or create a mobile tablet app for account managers that gives mobile visibility to current account data in our inventory management and payment processing systems.

5.3.4.3 Phase 3: Digital Scope

To address the company's approach to key areas of digital and outline the purpose, key initiatives and challenges of each objective.

This is where things start to get a little hairy (and potentially exhausting for some audiences). In company's digital scope, to take a deeper look at each objective which is outlined in company's Digital Base.

To put in company's Digital Scope, taking into account some insights on the most common objectives and channels, such as:

(a) Website(s). To have a branded website as it is a huge factor in buyers' buying decisions.
(b) Online Content. Blogging should be done strategically to produce astounding results like more leads, larger email lists and better search traffic.
(c) Digital Advertising/SEM/SEO. To target very specific users and the links actually go straight to the app store. As Facebook and Google have add programs specifically designed for mobile apps and potential customer are looking for:
    - credible sources
    - recommended by others
    - informative
    - relevant
    - visually educational
    - organized
    - appropriate
    - accompanied by a good user experience

To create blog posts and web pages that accomplish these things.

(a) **Customer Relationship Management (CRM)**. To customize the CRM use and to use the journey mapping example above by mapping out the entire sales process to ensure that the leveraging of CRM to the max.
(b) **Social**. To understand what role social business plays in your organization and sales funnel. The new path to success in Social is to integrate social into your other channels and customer experiences—both online and offline.

To create opportunities for company's customers to share.

5.3.4.4 Phase 4: Execution and Governance

Main aim: To prioritize the plan, taking into account which needs are the most urgent and important as well as current resources, timelines and budgets.

Prioritizing your projects. To have laser-like focus on just a handful of projects (usually 3–5) instead of overwhelming yourself with all of them.

To decide which projects go first, consider the impact each initiative will have on the following factors:

- Bottom-line
- Employee morale
- Productivity
- Brand Equity
- Customer Experience

To choose the handful of projects that will make the biggest impact and commit the time, money and resources to get them done.

Iterative Governance Models. To decide on a framework for how your digital project teams should function towards the end of the strategy phase. To empower the teams in order to make critical decisions as a high-functioning business unit, not a traditional department silo.

Finally, in summary, the way forward would be (Fig. 5.5):

5.3.5 Digital Transformation Strategy Development Plan

As observed in previous sections, all aspects (hereinafter the connecting elements) mentioned are connected; they superimpose on one other and not all are directly related to redefining and/or creating new technology

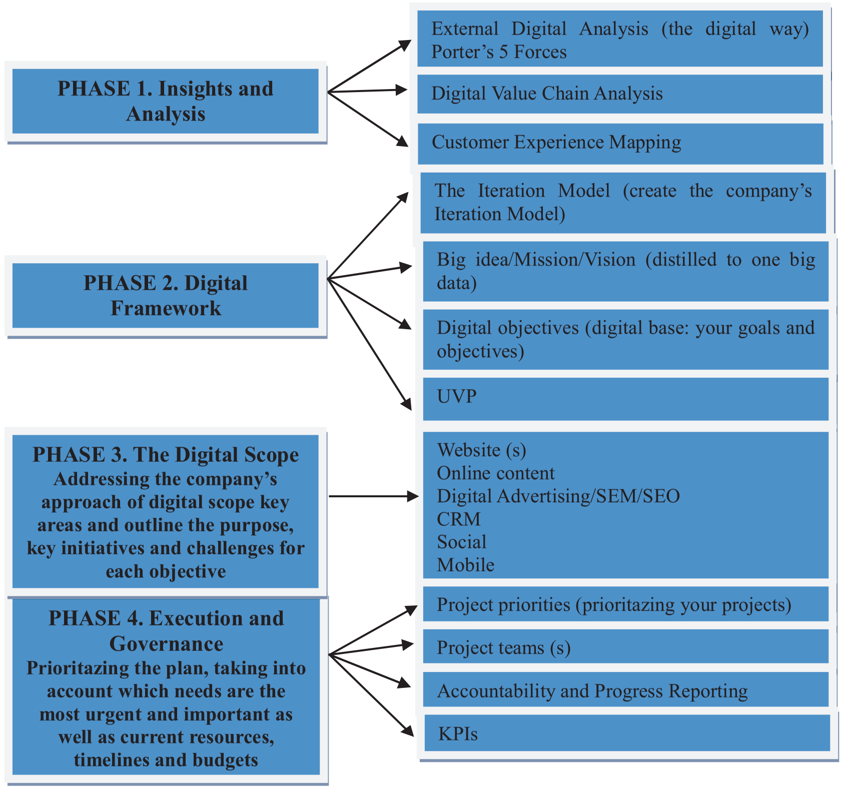

Fig. 5.5 Digital Transformation Strategy Framework (phases and actions/activities' interrelation) (Source: Author's own)

models. Although it is true that the main strategic focus is on the business; the definition of the Digital Strategy constitutes a comprehensive vision of digital transformation through which aspects such as customer experience, the evolution of technologies and innovation play a key role. But for the evolution of the development plan for the strategy, we must be aware that other elements interact and cannot be ignored.

The objective of the Digital Transformation Strategy is to create the capacities necessary for taking maximum advantage of the possibilities and opportunities offered by new technologies and their impact in a more rapid, better and more innovative way in the future.[2] A Digital Transformation journey needs a phase-based focus with a clear roadmap that involves a variety of interested parties, beyond the silos and limitations, internal and external. This roadmap (development plan) takes into account the fact the final objectives will continue to advance given that digital technology is a continuous journey, just like digital change and innovation are.

5.3.5.1 Digital Strategy Plan's Design

A Digital Strategy or plan is the articulation of an organization's vision, goals and purpose for engaging with digital solutions and technologies. It articulates the opportunities and challenges related to digital activities, the governance and management arrangements and risk management issues. It lays out a plan of action in order to maximize the business benefits of digital initiatives to the organization.

The connecting elements of the phases of development associated with the design of the development plan of the Digital Strategy will include those presented in Table 5.1.

5.4 AI Integration in Digital Transformation Strategy

Taking into account the above, we will explore the AI integration in the Digital Transformation Strategy, but first we will address the existence of a second framework, the "AI readiness framework".

[2] Definition available from https://www.i-scoop.eu/digital-transformation/

Table 5.1 The design of the Digital Transformation Strategy Development Plan phases

PHASE 1	**Analyse the external environment** (according to type of industry)	Adaptation to digital strategic framework	• Channel strategy • Assessment of opportunity	
	Analyse the internal environment **The client value chain**	Definition of level of digital maturity and diagnosis	Governance	
PHASE 2	**Define the business model**	• Assessment scorecard • Definition of digital strategy		
PHASE 3	**Define and analyse the target market**	PHASE 1 (analysis of external environment)	Type of industry	Specific characteristics of Industry 4.0 framework
PHASE 4	**Review and optimize**	Business and commercial process management		
		Business and automation processes	Adaptation to digital strategic framework	Operational processes
PHASE 5	**Improve the client experience**	Adaptation to digital strategic framework (client 360)		• Client knowledge • Client journey • Governance • "Active" listening
PHASE 6	**Foster a new organizational culture**	Adaptation to digital strategic framework	Operational processes	• Globalization • Skills development • Client-focussed culture • Innovation
PHASE 7	**Adopt the most appropriate technological solutions and perform continued benchmarking actions**	Adaptation to digital strategic framework	Technology and data	Governance
PHASE 8	**Know the new business ecosystems**			
PHASE 9	**Develop scorecards based on analytics to assess digital strategy**			

Source: Author's own

5.4.1 AI Readiness Framework

As J. Holmstrom (2021) asserts, strategies and means for selecting and implementing digital technologies that realize firms' goals in digital transformation have been extensively investigated. The recent surge in Artificial Intelligence (AI) technologies has amplified the need for such investigation, as they are being increasingly used in diverse organizational practices, creating not only new opportunities for digital transformation but also new challenges for managers of digital transformation processes.

A challenge for organizations adopting AI in their operations is that AI platforms vary in both scope and complexity, which hinders familiarity with them and hence their deployment to obtain competitive advantage (Iansiti & Lakhani, 2020). This springs partly from the "black box" nature of the algorithms (sets of digital instructions implemented to achieve defined goals) dictating AI responses, which are difficult to understand for members of organizations that are being increasingly shaped by AI (Hallinan & Striphas, 2016; Lindgren & Holmström, 2020; Pasquale, 2015). AI platforms are likely to transform organizations in qualitatively different ways from other technologies, so it is crucial to develop an understanding of organizations' abilities to meet these challenges (their AI readiness).

Since AI technologies have human-like cognitive capabilities (Huang & Rust, 2018), including knowing, learning, perceiving, sensing, acting, communicating and reasoning, their deployment may have far-reaching consequences for organizations and various associated ecosystem actors, including consumers, vendors, frontline service providers, and other stakeholders (Fernandes & Oliveira, 2021). However, there is a huge gap between the AI hype touted by AI vendors and its actual use in organizations. Thus, the importance of organizations' AI readiness in ongoing digital transformation has been recognized (Li et al., 2017; Pan, 2016). Moreover, there is a clear need for deeper exploration of AI's impact on organizational activities, boundaries and goals, including the mechanisms and processes involved in harnessing its power in digital transformation (Aldrich & Ruef, 2006).

Considering the above, we will explore AI integration in the Digital Transformation Strategy, but first we will address the existence of a second framework, the "AI readiness framework".

Therefore, J. Holmstrom (2021) presents a framework intended to assist efforts to address one of the first of these challenges: assessment of organizational AI readiness, that is, an organization's ability to deploy AI technologies to enable digital transformation, in four key dimensions: technologies, activities, boundaries and goals.

An AI readiness framework is a framework which invites fuller theorizing of the roles that AI can, and will, play in digital transformation.

For the purposes of this framework, he defines AI readiness as an organization's abilities to deploy and use AI in ways that add value to the organization. He presents a novel but an interesting approach in order to assess and visualize four key dimensions of this readiness: technologies, activities, boundaries and goals. The framework has been applied in workshop settings allowing an organization's members to grade their organization's capabilities (current and future potential) in each dimension. It highlights the risks and challenges involved in digital transformation associated with firm-wide mobilization of AI technologies for new ends, for which success rates are often low (e.g., Agarwal et al., 2010). The multidimensional AI readiness framework also provides a pragmatic tool that facilitates efforts to meet the associated challenges, reduce the risks and increase success rates.

5.4.2 AI Integration in Digital Transformation Strategy Model

According to Crossan et al. (1999), a good framework has three key requirements. First, it should identify the phenomenon of interest—in this case Digital Strategy and AI readiness. Second, the key premises or assumptions underlying the framework must be stated. A key premise underlying the presented Digital Strategy and AI readiness framework is that both are multi-dimensional, and that DT through AI affects all dimensions. Third, the relationships among elements of the framework must be described.

Therefore, the AI readiness framework[3] focuses attention on dimensions such as Technologies, Activities, Boundaries and Goals separately, but it is also intended to score AI Readiness as None, Low, Moderate, High and Excellent, and, taking into account two temporal perspectives, present and future, respectively.

[3] The duly detailed general approach available to interested readers is published in https://www.sciencedirect.com/science/article/pii/S0007681321000744

It is because of that, and according to J. Holmstrom (2021) approaches,[4] in the dimension "Technologies" it is examined at present: "Our present AI technology portfolio adds value to our organization", and in the future: "We have a strategy for using our AI technology portfolio to add value to our organization".

Secondly, in the dimension "Activities" it is examined at present: "Our present key activities are supported by AI in ways that add value to our organization", and in the future, "We have a strategy for using AI to support key activities in ways that add value to our organization".

Thirdly, in the dimension "Boundaries" it is examined at present: "Our present organizational boundaries are stretched by AI use in ways that add value to our organization" and in the future: "We have a strategy for using AI to change our organizational boundaries in ways that add value to our organization".

And finally, in the dimension "Goals" it is examined at present: "Our present AI use supports our goals in ways that add value to our organization" and in the future "We have a strategy for using AI to support our goals in ways that add value to our organization".

As we mentioned in Sect. 4.3 "AI in Business: Future Research Agenda" of Chap. 3 "The Impact of AI in Business, Economics and Innovation", and in order to guide the directions of future research addressing AI in business, Ashfaq et al. (2020) propose a set of research questions for certain business domains (Strategy, Relationship Marketing, Servicescape, Customer acceptance, Social acceptance, Management, Workforce and Transhumanism) that are still in need for further development. That is why we chose to include these business domains and research questions in the main elements of the DT Strategic Framework and the AI readiness framework.

Having said this and taking into account AI in business: future research elements, framework dimensions of AI readiness and the main elements of the DT Strategic Framework, we present a graphical representation of the proposed "AI integration in the Digital Transformation Strategy Model" in Fig. 5.6.

[4] *Appendix 2. Questionnaire for the AI readiness framework and scores awarded for the insurance firm.*

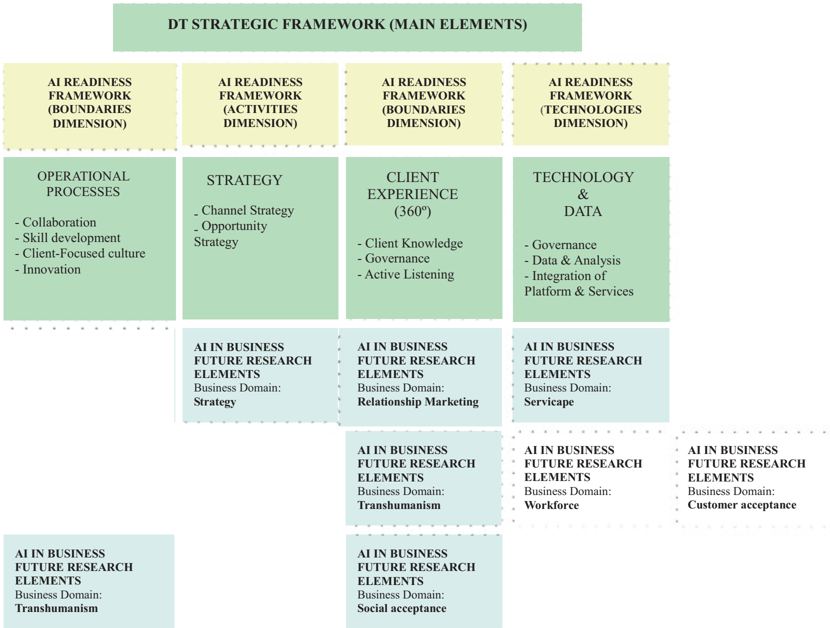

Fig. 5.6 AI integration in Digital Transformation Strategy Model. (Source: Author's own)

Where, as seen previously, AI technology integration (mechanisms and instruments) would be conveniently reflected and rooted within the digital business transformation strategy.

References

Adams, R., Bessant, J., & Phelps, R. (2006). Innovation Management Measurement: A Review. *International Journal of Management Reviews, 8*(1), 21–47. Retrieved November 5, 2016, from http://oapen.org/search?identifier=560292

Agarwal, R., Gao, G., DesRoches, C., & Jha, A. K. (2010). Research Commentary—The Digital Transformation of Healthcare: Current Status and the Road Ahead. *Information Systems Research, 21*(4), 796–809.

Aldrich, H. E., & Ruef, M. (2006). *Organizations Evolving*. SAGE Publications.

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics, 54*, 101473.

Bergeron, F., Raymond, L., & Rivard, S. (2004). Ideal patterns of strategic alignment and business performance. *Information & management, 41*(8), 1003–1020.

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly, 37*(2), 471–482.

Chesbrough, H., & Rosenbloom, S. R. (2002). The Role of the Business Model in Capturing Value from Innovation: Evidence from Xerox Corporation's Technology Spin-offs Companies. *Industrial and Corporate Change, 11*(3), 529–555.

Crossan, M. M., Lane, H. W., & White, R. E. (1999). An Organizational Learning Framework: From Intuition to Institution. *Academy of Management Review, 24*(3), 522–537.

Descartes, R. (1983). *Principles of Philosophy* (V. R. Miller & R. P. Miller, Trans.). Springer Netherlands.

Descartes, R. (1998). Descartes' Meditations: background source materials. Cambridge University Press.

Echeberria, L. (2020). *A Digital Framework for Industry 4.0*. Springer International Publishing.

Engel, C. (1999). *The Internet and the Nation State*. Lectiones Jenenses, 21.

Fernandes, T., & Oliveira, E. (2021). Understanding Consumers' Acceptance of Automated Technologies in Service Encounters: Drivers of Digital Voice Assistants Adoption. *Journal of Business Research, 122*, 180–191.

Hallinan, B., & Striphas, T. (2016). Recommended for You: The Netflix Prize and the Production of Algorithmic Culture. *New Media & Society, 18*(1), 117–137.

Harmancioglu, N., Droge, C., & Calantone, R. J. (2009). Strategic Fit to Resources versus NPD Execution Proficiencies: What are Their Roles in Determining Success? *Journal of the Academy of Marketing Science, 37*(3), 266–282.

Harrison, J. (2015). *The Digital Strategy Guide*. Retrieved August 2018, from http://www.harrisonjlloyd.com/2015/03/30/the-digital-strategy-guide-i-couldnt-find-with-examples-and-templates/#structure

Holmstrom, J. (2021). From AI to Digital Transformation: The AI Readiness Framework. *Business Horizons*. https://doi.org/10.1016/j.bushor.2021.03.006

Holmstrom, J., & Nyle, D. (2015). Digital Innovation Strategy: A Framework for Diagnosing and Improving Digital Product and Service Innovation. *Business Horizons, 58*(1), 57–67.

Huang, M. H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research, 21*(2), 155–172.

Iansiti, M., & Lakhani, K. R. (2020). *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Harvard Business Press.

Li, B. H., Hou, B. C., Yu, W. T., Lu, X. B., & Yang, C. W. (2017). Applications of Artificial Intelligence in Intelligent Manufacturing: A Review. *Frontiers of Information Technology & Electronic Engineering, 18*(1), 86–96.

Lindgren, S., & Holmström, J. (2020). A Social Science Perspective on Artificial Intelligence: Building Blocks for a Research Agenda. *Journal of Digital Social Research, 2*(3), 1–15.

Pan, Y. (2016). Heading Toward Artificial Intelligence 2.0. *Engineering, 2*(4), 409–413.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Porter, M. E. (2001). Strategy and the Internet. *Harvard Business Review, 79*(2), 62–78.

Powell, T. (1992). Organizational Alignment as Competitive Advantage. *Strategic Management Journal, 13*, 119–134.

Rhem, A. J. (2021). AI Ethics and Its Impact on Knowledge Management. *AI and Ethics, 1*(1), 33–37.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long range planning, 43*(2–3), 172–194.

Tidd, J. (2001). Innovation management in context: environment, organization and performance. *International journal of management reviews, 3*(3), 169–183.

Tidd, J. (2006). A review of innovation models. Imperial College London, 16.

VanBoskirk, S., Gill, M., Green, D., Berman, A., Swire, J., & Birrell, R. (2017). The digital maturity model 5.0. *Forrester Research*.

Zott, C., & Amit, R. (2013). The business model: A theoretically anchored robust construct for strategic analysis. *Strategic organization, 11*(4), 403–411.

CHAPTER 6

Case Studies of Real AI Applications

José Antonio Álvarez López

Abstract Companies are obtaining competitive advantages by using AI and automated machine learning systems.

This chapter analyses the best practices of 11 broad sectors that most employ these technologies to improve their efficiency, personalise their services and increase their growths. The chapter presents several 'Cases of AI Use and Applications in every industry and business functions; putting artificial intelligence to work', evaluating and implementing business applications: Customer service, Consumer: Marketing and sales, Energy, resources and industries, Financial services and FinTech, Government and Public Administration, Life Sciences and healthcare, HealthTech, Retail, Mobility and self-driving cars, Human Resources, Operations and Startups.

Keywords Artificial intelligence • Automated machine learning systems • Advanced sectors • Companies • Robotic process automation • Deep learning

J. A. Álvarez López (✉)
Universidad a Distancia de Madrid, Madrid, Spain
e-mail: jalvarez@exes.es

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_6

6.1 Putting Artificial Intelligence to Work: Evaluating and Implementing Business Applications

The New Economy makes markets more competitive, with customers having progressively greater needs and wishes with respect to the immediacy of services and products. Further, markets are global and companies operate in a multitude of markets, industries and sectors. In this context, companies face enormous challenges and must innovate constantly, to be more efficient, to have new services and products available and to be better. Technology in general and, in particular, machine learning systems and artificial intelligence (AI) can facilitate these change processes, creating extremely solid and sustainable competitive advantages.

Some of these innovations will be described next, broken down by sectors. We will show how some industries are capable of obtaining extremely important improvements in efficiency by employing these technologies.

6.2 AI Use Cases and Applications for Every Industry and Business Function

This section describes different extremely successful use cases, with the information organised into 11 sectors with intensive use of AI and machine learning systems. To facilitate reading, we have followed a common structure with five sections: Title, General overview, Findings, Implementation and Recommendations for companies.

6.2.1 Customer Service

Title: Advanced customer service systems with AI and chatbots

General Overview:

Increasing customer satisfaction is an obligatory objective in all companies' strategies. There are a multitude of AI-supported systems, including chatbots (Dilmegani, 2021a) that can greatly assist with increasing satisfaction.

Findings:

Systems like *social listening* and *ticketing* are used, taking advantage of natural language processing to identify customers, respond to them and redirect them to the right department. *Intelligent call routing* assigns and routes calls to the right agent. *Call classification* detects customers' needs with natural language processing. *Voice authentication* can authenticate users without a need for passwords. *Chatbots* are employed to understand the more complex queries customers make and provide them with more precise and exact service, available 24×7 (Dilmegani, 2021b).

Implementation:

The types of systems to implement include a need to use advanced processes, such as natural language voice recognition, although there are market solutions available that can be used, like those offered by *Google*. These systems tend to be very profitable, with investments recouped quickly. For example, ING Bank observed a 15% increase in the quality indicator for its sales and a 3% drop in silence rates, after integrating AI into its call systems (Sestek, 2020).

Recommendations for Companies:

Customers increasingly demand better service at any time, immediately and want their problems resolved, which is why systems with AI will have a growing prominence at businesses. Learning from their implementation can be a competitive future in the near future.

6.2.2 Consumers: Marketing and Sales

Title: Sales management and customer service systems with automation and AI

General Overview:

Automated systems with AI are being used as an evolution of customer relationship management (CRM), for example, managing data collection, launching personalised products and services and prescriptive selling (Dilmegani, 2021a).

Findings:

AI is used for the internal management of sales, employing—for example—*sales data input automation*, in which data from several sources are transferred automatically to the CRM system, unified and effort-free. The data obtained are also analysed to match customers' wishes for products, via *sales content personalisation and analytics.* AI systems are even employed to improve salespeople's processes, identifying prior successful patterns, with *Sales Rep Next-Best-Action* suggestions. Finally, interactions with customers are analysed to predict their purchases, by comparing them to similar customers with *prescriptive sales* processes.

Implementation:

These systems are an evolution from CRM, whose implementation is extremely widespread at companies. Investment in these systems is easily recouped, due to being directly related to increasing sales. Some examples are given below (Daley, 2021).

AMPLERO: building customer relationships. *Amplero* creates AI-driven marketing tools for a wide range of industries focused on consumers, including finance, retail telecommunications and games.

Amplero algorithms detect patterns in data to build dynamic target market profiles. Marketing and sales specialists can use the company's machine learning software and capacities to execute thousands of experiments at scale, which let acquisition costs be reduced and increase average revenue per customer. Its AI systems are used by important companies, including Sprint, Microsoft and TaxAct.

DRIFT: conversational marketing. Drift uses chatbots, machine learning and natural language processing to help companies make their sales cycle more efficient and arrange meetings.

Technology is particularly good at automating marketing tasks that traditionally ate up a lot of time. For example, after a customer reaches a website using *Drift*, a chatbot appears and asks questions that identify whether or not it is really a potential customer and routes the person correctly. The company's *Drift Assistant* also automates email responses, the routing of potential customers and the updating of contact details. Drift is currently used by over 100,000 companies.

Recommendations for Companies:

CRM systems are always recommendable for optimising sales in any sector, particularly those with highly fragmented customers. By considering repetition patterns of successful sales and highlighting best practices, sales optimisation is facilitated.

6.2.3 Energy, Resources and Industries

Title: AI and robotics to improve energy production, natural resources and for industry

General Overview:

Industry is one of the activity sectors with the highest implementation of AI. Robotics is revolutionising industry and obtaining resources (Joyanes Aguilar, 2018; Daley, 2021). Sensors collect progressively more information that feeds AI systems, obtaining incremental efficiency improvements in these sectors. Uses are also considered to fight climate change (Hao, 2019).

Findings:

AI is used in different sectors of industry, and is a huge support to ensure that industrial processes generate greater performance wherever it is used. It can independently handle different processes, in coordination with other agents. For example, AI is used in manufacturing processes to identify possible system failures and get ahead of events, monitoring them via quality systems with computer vision, neuron networks and pattern recognition. They also detect imperfections in industrial processes, for example parts that are in bad condition, so that they can be repaired.

Through drones and photogrammetry, land areas can be studied to determine—for example—the best site for crops, precision agriculture, the existence of gas pockets, for energy production, or on fires in certain places.

Further, some applications of AI in these sectors can be listed, which can combat climate change (Hao, 2019): predicting how much energy will be needed, discovery of new materials, route optimisation, reducing barriers for electric vehicles, energy efficiency of buildings, supply chain optimisation, precision agriculture at scale, monitoring deforestation and so on.

Implementation:

They are generally expensive systems and their commissioning requires highly experienced professionals. Nonetheless, given the amortisation periods of industry, they could turn out to be very interesting. Conversely, they are very cheap for agriculture and yields are very attractive.

Recommendations for Companies:

Industry has been implementing these systems for many years and continuous improvement is leading to progressively more efficient companies. In agriculture and energy production, AI systems are being deployed very quickly, as there are huge opportunities in these sectors.

6.2.4 Financial Services and FinTech

Title: AI adapted to financial services and FinTech

General Overview:

Historically, financial services are one of the sectors that most intensively uses technology, so it should come as no surprise that they are also pioneers in applying AI-based technologies. The financial industry has many customers, a lot of data and enormous growth possibilities, meaning that it also has an enormous opportunity for incorporating AI and machine learning. It has also developed a new sector that joins technology and financial services—FinTech—which facilitates the incorporation of extremely innovative solutions to industry.

Findings:

AI in the financial sector is present in scores of processes (Dilmegani, 2021a). They include billing, automated services that remember customers' bill payments; robo-advisors, mobile app assistants and chatbots to control personal finances; regulatory compliance, which uses natural processing language to distribute legal and regulatory texts, handling thousands of procedures with no need for human interaction; data gathering employs AI to collect external data on the market; debt collection, which uses AI in debt collection processes and conflict resolution. The development of other solutions at some FinTech companies is relevant, which use AI for fraud detection, which reduces the operating costs greatly for this activity. AI is also used—for example—in detecting risk in insurance quotes and banks' ratings of customers. There is enormous development on financial analysis platforms to establish investment strategies.

Implementation:

Industry is aware of the precision and efficiency of AI-driven processes and there are many examples of success (Daley, 2021). To mention just a few:

Betterment: Robo-advisor pioneer. It uses AI on an automated financial investment platform with robo-advisor technology, which learns about investors and builds their personalised profiles based on their financial plans. Algorithms are used to collect data, tax information, trading, portfolio management and so forth. These tasks previously required large-scale human interaction. At present, the company manages 10,000 million dollars and provides services to 250,000 customers.

Another example is *Alphasense*, which has a financial search engine with AI, which helps customers obtain informational advantages. It combines a linguistic search and natural language processing to analyse the key data of 35,000 financial institutions. According to the company, 800 firms use its services.

Recommendations for Companies:

The financial industry is extremely competitive and the application of AI entails an enormous advantage that some companies are already using. This sector will undoubtedly go through restructuring in upcoming years, during which AI will play a leading role.

6.2.5 Government and Public Administration

Title: AI applied to public administrations and their millions of users

General Overview:

Public administrations are a very favourable environment for the application of machine learning and AI (Blat, 2019), due to having millions of users and these users having a multitude of interactions with the administration. Further, many of these interactions are recurrent, making it possible to implement continuous learning systems.

Findings:

There are scores of experiences where AI is applied at the public administration (Campos, 2017). For example, with conversational chatbots, to assist with 24×7 customer service, adapted to people's different profiles. Border surveillance systems and roads with drones equipped with cameras with facial or number plate recognition. The application of AI and Big Data to taxation and collection systems to—for example—optimise taxes or detect fraud. In any event, the implementation level of these types of systems must go hand-in-hand with a deep change in the administrations, supported by digitalisation, especially of their professionals.

Implementation:

There are a multitude of systems implemented that could use AI, tracing and detection systems of infected people in the recent coronavirus pandemic is one example that could also streamline and speed up vaccine appointments, by natural language processing among the older age groups. Tax and fee verification systems can also be implemented, once again, via natural language processing.

Recommendations for Companies:

In upcoming years, the public administrations will be one of the widespread implementers of AI and they should clearly learn from the best possible practices. Companies will find an enormous opportunity in this process, particularly those directly related to this industry, engineering, IT and so on.

6.2.6 Life Sciences and Healthcare, HealthTech

Title: Life sciences and healthcare revolutionised by using AI

General Overview:

AI is transforming healthcare in practically all areas of this industry (Dilmegani, 2021a). There are numerous examples where AI rationalises processes and substantially improves service efficiency and quality. Further, the application of these technologies is opening up new and innovative solutions that will undoubtedly lengthen many people's life expectancy and quality of life. Once again, the setting is ideal, millions of users, with similar needs and recurrence that in upcoming years will definitely see enormous advances in sectors related to health, medical care and disease prevention. A technology sector highly specialised in health is also being developed—HealthTech—that presents very innovative solutions.

Findings:

There are many success stories of AI applied to healthcare (Daley, 2021). Some of them are related to changes in medical care, which concentrates a significant part of the costs due to their inefficient processes. Further examples include achieving a reduction in the number of unnecessary visits that patients make, freeing up to 20% of nurses' time; assistants in work flows, which will free up some 17% of doctors' schedules; pharmaceutical companies reducing the research timelines into new drugs; robot-assisted surgeries improve precision during operations and reduce hospitalisation times; protection against cybercriminals with regard to personal medical data and so on. Diagnoses are also being improved by analysing patients' data with machine learning systems.

Implementation:

Given the huge size of the industry, the vast possibilities offered by AI and the primordial importance of health in all economies mean that there are many ground-breaking initiatives in this field. To mention a few:

PathAI: improving diagnostic pathology. Its machine learning algorithms help pathologists to analyse tissue samples and to more precisely diagnose them to then improve treatments. *PathAI* has worked with the *Bill y Melinda Gates Foundation* and with the company *Phillips* to develop mass-access diagnostic tests.

PAGER: proactive healthcare management. The company analyses medical records with AI to find patients not covered by treatments, which is a service similar to an assistant who helps with these tasks, as well as helping patients book appointments and make payments. The application enables chatting 24×7 with a nurse or a doctor, and even writing prescriptions.

ATOMWISE: streamlining drug discovery. It uses AI and deep learning to facilitate the discovery of new drugs. Via neuron networks, the company can '*extract insights from millions of experimental affinity measures and thousands of protein structures to predict the binding of small molecules* *to proteins*' (Atomwise, 2018). It has improved success rates by 10,000 times and examines 10–20 million compounds a day by using AI. Its technology is accelerating the work of clinical analyses. *Atomwise* is being employed to handle Ebola and multiple sclerosis.

Massachusetts General Hospital: AI in the ER and beyond. Massachusetts General Hospital has formed a partnership with the IT company *NVIDIA* to implement AI-driven machines in processes to detect, diagnose, treat and manage diseases.

The programs are fed by over 10,000 million medical radiology and pathology images to facilitate the speed of tests and, therefore, diagnostic capacities.

The hospital concluded a pilot system that uses AI for the rapid preselection of patients for pneumothorax, more commonly known as 'collapsed lung'. The results were promising enough to lead to plans to implement the technology in emergencies.

There are also other extremely interesting initiatives (Daley, 2020).

The use of AI to reduce medical errors, more precise cancer diagnoses with individualised medical treatment methods (implemented by *PathAI*), a smart symptoms verifier (implemented by *Boya*) that, via a chatbot, listens to customers' health problems to guide them to the right care level depending on their pathology.

In short, AI simplifies the lives of patients, doctors and hospital administrative staff by performing tasks normally done by humans, but in less time and at a fraction of the cost.

Recommendations for Companies:

The healthcare industry is enormous both in size and possibilities, and is one sector that will grow the most in upcoming years, as health is one of citizens' main concerns. In this context, AI and its applications will progressively be implemented exponentially, creating opportunities at the best companies offering these types of services.

6.2.7 Retail

Title: Retail and e-commerce improved with AI

General Overview:

Retail sectors, given their enormous volume of customers, products and transactions, were one of the first sectors to incorporate machine learning technologies. AI and its algorithms are implemented in e-commerce processes, building more personal relationships with customers, adapting the experience to each user, increasing sales and facilitating relationships with greater loyalty that are longer lasting.

Findings:

In retail, AI is employed via chatbots, to predict purchases, compile data and thus create a more customer-centric e-commerce experience. Indeed, AI systems are being used to increase customer loyalty and, in this way, increase sales.

Implementation:

The e-commerce and retail industry is extraordinarily competitive, as it is no easy feat to stand out either by product or price, so that many market players are implementing systems that improve service and the user experience. Below, some of the most successful companies are listed (Daley, 2021).

Amazon. It has implemented AI across the boards. *Amazon* is one of the main worldwide e-commerce players and supports a multitude of processes in AI, for example, purchase recommendations, warehouse and inventory robots that collect, classify and send products, to name just a couple. It has Alexa, its voice assistant that operates on millions of devices. *Amazon* is one of the outstanding examples of AI application, as it is present in almost all of its processes.

Twiggle. NLP for e-commerce search. *Twiggle* incorporates an advanced search engine for e-commerce with natural language processing. The combination of deep learning and retail know-how combine to offer customers exactly what they need. *Twiggle* confirms that a site with two million monthly visits could lose up to 266,600 customers due to bad searches. For customers using its search system, the company can increase its 'add to cart' by up to 9% and increase its click-through rate by 12%.

Another use of AI in retail is the improvement of internal processes (Lenovo, 2021). Adding these systems can reduce delays and operating costs, increase the security of customers' data, reduce human failures in order delivery, make processes more agile and so forth. It is also effective at generating reports, which creates more control in process management, process automation, improvements to infrastructures and many more.

AI is employed in other retail areas (Aarón, 2019), in features such as predictive analysis in inventory management, in pricing and promotional strategies, in optimising recommendation engines, and many more.

Recommendations for Companies:

In summary, AI solutions in the retail sector are generally very well received by industry, due to being directly related with increased sales, obtaining highly visible returns on investment.

6.2.8 Mobility and Self-Driving Cars

Title: AI revolutionises transport by developing self-driving cars

General Overview:

The transport and mobility sector represents 5.6% of the GDP in OECD member countries (OECD, 2018), meaning it is a sector with enormous importance with regard to size. Further, self-driving cars are expected to have an enormous impact in—for example—savings through accident reduction, fewer traffic jams and reducing pollution (OECD, 2019). Estimates claim that 1100 lives could be saved and 38,000 million dollars could be saved in the United States via the 10% adoption of self-driving cars. An adoption of 90% could improve these numbers even more, preventing 21,700 deaths and saving 447,000 million dollars in the USA (Fagnant & Kockelman, 2015). Therefore, this sector has great importance, where AI advances are improving the adoption levels, achieving increasingly safer vehicles.

Findings:

The more disruptive use of AI in vehicles is aimed at the six standard levels of driving automation (ORAD, 2016). They are also working on cybersecurity protection of vehicles with AI, like in other industries. Advances are also being made in vehicle vision systems, to detect thousands of data to tenths of seconds, people, signals, other vehicles and so forth, executed with AI learning systems.

Implementation:

There are numerous initiatives in this sector, which employs AI as their main technology, providing progressively more advanced driving assistants for vehicles. Today, it is common for cars to park automatically and safely manage braking systems, such as ABS. A few examples are given below (Daley, 2021).

CRUISE: self-driving into the future. Cruise is at the vanguard of self-driving cars and its vehicles are some of the first in the world to hit the streets, using AI. Its autonomous vehicles collect a petabyte of information every day. AI uses these data to improve safety, driving techniques, the most effective routes and so forth. The company has an electric driverless shuttle—the *Cruise Origin*—which has travelled over one million miles in San Francisco autonomously and with no accidents.

WAYMO: an autonomous fleet of vehicles. Waymo is the autonomous vehicle project by Google, which has designed a range of driverless vehicles to meet the needs of drivers throughout the country, private users, ride-sharing and lorry companies.

Waymo vehicles have already travelled millions of miles through over ten states, using AI on the road to collect and analyse data. They use an advanced sensor system and AI predicts what could happen. Thanks to AI, Waymo vehicles can resolve driving situations, without humans having to interact at the wheel.

Waymo One, the company's transport service is currently being used in the Phoenix region to pick up passengers and take them to their destinations, all autonomously, of course. Meanwhile, Waymo Via centres on autonomous vehicles for long-distance logistics and last-mile delivery.

LUMINAR TECHNOLOGIES: seeing the future. Luminar Technologies produces vision systems based on the most advanced LIDAR in the world. These systems collect information with laser, which can perceive the depth of what is ahead. Technology ensures that artificial-intelligence-based software programs see people, objects, events, road conditions and many other features from over 250 metres away, so that autonomous vehicles are given a long time to analyse and react to any situation that occurs. The innovative technology has the capacity to measures the speeds of different objects, so that an AV can easily determine optimal movement in very little time, meaning that they are extremely safe systems.

Recommendations for Companies:

The AI industry applied to mobility for self-driving cars has an enormously promising future and there are increasingly more applications that are improving the safety of these vehicles. New business models are also being implemented that are based on ride-sharing and many important companies are incorporating them into their service catalogues, such as *Access by BMW*, the *Mercedes Collection* and the *Porsche Passport*.

In summary, there are clear advances being made in these technologies that will depend, on the one hand, on innovation capacity and, on the other, on suitable regulations that will permit their development.

6.2.9 HR

Title: AI improves people and talent management

General Overview:

Managing people is one of the main concerns of today's companies, at a time when talent is so scarce. For this reason, AI systems have been developed to improve the value of companies' employees.

Findings:

AI is being employed in at least four issues related to people management (Dilmegani, 2021a): 'hiring and recruitment' (Dilmegani, 2021c) to identify the most valuable candidates; 'performance management' to fairly evaluate employees; 'retention management' to detect those employees who could leave the company and the reasons for it and 'human resource analysis' (Dilmegani, 2021d) to analytically study the status of employees and act consequently.

Implementation:

There is an ecosystem of vertical initiatives for different fields related to human resources. For example, in recruitment issues, where the difficulty rests with finding the right talent, selection teams have to spend a significant amount of their time on automatable tasks such as selecting CVs, but with AI machine learning can be employed to reduce time spent in this area. Concretely, AI and machine learning are used to study microdata related to the candidates' skills and capabilities.

Recommendations for Companies:

Today's people-centric companies are always interested in having the best talent at their organisations and this will never end. For this reason, all tools aimed at improving it will be welcome. There are scalability options in this sector that are extremely attractive to innovative companies.

6.2.10 Operations

Title: Innovations in operations: Robotic process automation (RPA) and AI

General Overview:

Operational efficiency is one of the most important divisions at many companies, and affects a multitude of areas including logistics, manufacturing and sales. For this reason, there has been great preoccupation with improving it. Technology has not been outside this need and digitalisation and robotisation processes have been incorporated, called RPA, which is also currently supported by AI.

Findings:

According to a report by McKinsey (Lhuer, 2016), companies that incorporate RPA have a potential ROI between 30% and 200% the first year. Further, digitalisation with these systems does not require replacing legacy systems, where they learn how to use them by being integrated with them. This has led to widespread deployment of these solutions in many industries.

Implementation:

There is enough experience with implementing RPA systems and incorporating AI is one more step in improving them. Specifically, AI is used within RPAs for many processes, where prominent uses include: process mining, which uses algorithms to understand real processes in detail; predictive maintenance to minimise interruptions; manufacturing analysis, which analyses the complete industrial process to increase efficiency; inventory and supply chain optimisation; robotics, exchanging collaborative bots to interact better with employees; cash-free payments; and billing.

Recommendations for Companies:

There are so many possibilities that can be implemented in this company division and the market is so competitive that all organisations should consider incorporating RPA systems, opening up enormous opportunities for the companies that provide these services.

6.2.11 Startups

Title: Startups clearly commit to AI for their innovative projects

General Overview:

When they began, startups were closely linked to innovation and technology, and even to disruptive technology, understood as a permanent innovation process that makes what came before obsolete. Today, the success of a startup is especially connected to the good use of resources. For example, at startups where work consists of repetitive tasks, AI can take on a human's workload with no problem (TodoStartups, 2020).

Findings:

There are numerous examples of AI-based applications that came from startups, some of which are published in the CB Insights report (CB Insights, 2021), in its annual list of the 100 most promising private AI companies in the world, obtained from among over 6000 candidates. The most successful trends include applications for industrial AI in a multitude of sectors, such as retail, legal and gaming. There are also examples of cross-cutting solutions, seeking its use in different sectors, but with common needs in areas of corporate and operational intelligence. Other examples can be found in computer tools for data processing and software for automating and simplifying IT operations.

Implementation:

The commissioning of AI applications is not easy, but it is essential for the market to adopt them so that innovation is complete. To evaluate the implementation level of solutions, different factors can be considered, as CB Insights points out. Different issues can be studied, such as patent analysis, the health of the startup, financial info, the company's traction, commercial relations, the size of the market, the competition, relevance and so forth.

Recommendations for Companies:

When speaking of the startup sector, it is usually not simple to determine which solution will be the first on the market, but the fact that they are many whose developments are supported by AI is a very positive indicator, making it an extremely attractive opportunity to bet on these types of companies with such advanced technologies.

6.3 Conclusions

As the outcome of what we have set out above and what was studied in the preliminary analysis, we present these conclusions:

- AI-driven technologies and machine learning are advantageous for the majority of industries, and there are both great interest and investment volume, which has led to the existence of a multitude of successful initiatives.
- There are solutions implemented in industry that obtain extraordinary returns on investment, in efficiency improvements, in resource optimisation and in personalising services.
- AI solutions are particularly interesting in markets with highly fragmented customers, with a large quantity of data and recurrence in service needs, like retail, healthcare services and public administrations. These solutions are also very efficient at improving processes, such as robotisation in manufacturing processes.
- There are different speeds in adopting AI solutions by sectors, where industrial sectors have a large scope compared to others, which demand these solutions less.
- There is an enormous opportunity in adopting AI. On the one hand, by part of the companies in industry, which can obtain enormous competitive advantages and, on the other, by companies that provide this type of service, due to being a market still undergoing growth.
- The use of AI in all sectors will grow exponentially in upcoming years, creating solutions that we cannot even imagine still.

References

Aaron. (2019). The top 5 popular use cases of AI in retail – Chain of demand. Retrieved July 12, 2021, from https://www.chainofdemand.co/the-top-5-popular-use-cases-of-ai-in-retail/

Atomwise. (2018). Atomwise – Better medicines faster. Retrieved July 12, 2021, from https://www.atomwise.com/

Blat, F. (2019). Inteligencia Artificial en la administración pública: casos de uso Gobierto. Retrieved July 11, 2021, from https://gobierto.es/blog/20190920-inteligencia-artificial-administracion-publica.html

Campos, C. (2017). 5 ejemplos de Inteligencia Artificial (IA) en la Administración pública ¿Presente o futuro? – Concepción Campos. Retrieved 11 July 2021 from https://concepcioncampos.org/5-ejemplos-de-inteligencia-artificial-ia-en-la-administracion-publica-presente-o-futuro/

CB Insights. (2021). AI 100: The artificial intelligence startups redefining industries. *CB Insights*. https://www.cbinsights.com/research/report/artificial-intelligence-top-startups/

Daley, S. (2020). 32 AI in healthcare: Examples you should know | Built In. Retrieved July 12, 2021, from https://builtin.com/artificial-intelligence/artificial-intelligence-healthcare

Daley, S. (2021). 23 top artificial intelligence examples you should know 2021 | Built In. Retrieved July 11, 2021, from https://builtin.com/artificial-intelligence/examples-ai-in-industry

Dilmegani, C. (2021a). 100+ AI use cases & applications in 2021: In-depth guide. Retrieved July 11, 2021, from https://research.aimultiple.com/ai-usecases/#customer-service

Dilmegani, C. (2021b). Chatbots: In-depth conversational bots guide [2021 update]. Retrieved July 11, 2021, from https://research.aimultiple.com/chatbot/

Dilmegani, C. (2021c). Recruiting AI in 2021: Guide to augmenting the hiring team. Retrieved July 13, 2021, from https://research.aimultiple.com/recruiting-ai/

Dilmegani, C. (2021d). 6 reasons to set up HR analytics [2021 update]. Retrieved July 13, 2021, from https://research.aimultiple.com/benefits-of-hr-analytics/

Fagnant, D., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181. https://www.sciencedirect.com/science/article/pii/S0

Hao, K. (2019). Here are 10 ways AI could help fight climate change. *MIT Technology Review*. https://www.technologyreview.com/2019/06/20/134864/ai-climate-change-machine-learning/

Joyanes Aguilar, L. (2018). *Industria 4.0: la cuarta revolución industrial*. Marcombo.

Lenovo. (2021). 7 ejemplos del uso de la inteligencia artificial en las empresas. *Lenovo Tech Today Mexico*. Retrieved July 12, 2021, from https://techtoday.lenovo.com/mx/es/solutions/smb/inteligencia-artificial-empresas

Lhuer, X. (2016). The next acronym you need to know about: RPA (robotic process automation). *Digital McKinsey, 17*, 1–5. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-next-acronym-you-need-to-know-about-rpa#

OECD. (2018). Structural Analysis Database (STAN), Rev. 4, Divisions 49 to 53, Retrieved July 13, 2021, from http://www.oecd.org/industry/ind/stanstructuralanalysisdatabase.htm

OECD. (2019). *Artificial intelligence in society*. OECD Publishing. https://doi.org/10.1787/eedfee77-en

ORAD. (2016). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. On-Road Automated Driving (ORAD) Committee, SAE International. https://doi.org/10.4271/j3016_201609

Sestek. (2020). ING Bank call center speech analytics case study | *Sestek*. Retrieved July 11, 2021, from https://www.sestek.com/case-studies/ing-bank-speech-analytics-case-study/

TodoStartups. (2020). Startups y el uso de inteligencia artificial – *Todostartups*. Retrieved July 13, 2021, from https://www.todostartups.com/nuevas-tecnologias/startups-y-el-uso-de-inteligencia-artificial?cn-reloaded=1

CHAPTER 7

Technological Singularity and Ethical Issues of AI

David Lizcano Casas and Juan Pazos Sierra

Abstract The technological singularity is the possibility of achieving an artificial general intelligence that surpasses human intelligence, one of the major paradigms of life for humanity today. Nevertheless, there are only opinions on whether this goal will be achieved or not, without any concrete studies that quantify the possibilities of this event taking place or how it might take place. This chapter delves into these questions applied to the area of business.

Keywords Singularity • Paradigm • Business • Ethical

7.1 Introduction

In 1847, Richard Thornton wrote the following lines on the recent invention of a mechanical calculator with four basic functions: "that such machines, by which the scholar may, by turning a crank, grind out the solution of a problem without the fatigue of mental application, … But

D. L. Casas (✉) • J. P. Sierra
Universidad a Distancia de Madrid (UDIMA), Madrid, Spain
e-mail: david.lizcano@udima.es; juan.pazos@udima.es

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9_7

who knows that such machines when brought to greater perfection, may not 'think' to a plan to remedy all their own defects and then grind out ideas beyond the ken of mortal mind!" (Thornton, 1847).

Subsequently, Samuel Butler wrote under the pen name of "Cellarius": Reflect upon the extraordinary advance which machines have made during the last few hundred years, and note how slowly the animal and vegetable kingdoms are advancing. The more highly organised machines are creatures not so much of yesterday, as of the last five minutes, so to speak, in comparison with past time. Assume for the sake of argument that conscious beings have existed for some 20 million years: See what strides machines have made in the last thousand! May not the world last 20 million years longer? If so, what will they not in the end become? And further on he write: We cannot calculate on any corresponding advance in man's intellectual or physical powers which shall be a set-off against the far greater development which seems in store for the machines (Cellarius, 1863). As a matter of fact, Butler included this essay in Chapters 23–25 of his Erewhon (Butler, 1872–1901).

In 1909, Henry Adams, the developer of a "physical theory of history" proposed a law of the acceleration of thought by applying Newton's law of gravity. In it, he interpreted history as a process of advancing towards balance, speculating that said process would take thought to the limits of its possibilities in the year 1921, adding that it was highly possible. And he added that the consequences could be as surprising as water transforming into steam, the worm into the butterfly, from electrons to the radio (Adams, 1920).

7.2 Turing and the Technological Singularity

And thus we arrive at the year 1951 when Turing spoke of machines intellectually surpassing human beings, at a conference held in Manchester: "once the machine's thinking method has begun, it will not take long to overcome our weak problems. … Therefore, at a certain moment, it must be expected that machines would take control in the manner mentioned in Samuel Butler's Erewhon."

Years later, in 1965, Turing's colleague at Bletchley Park, Irwing John Good, wrote about an explosion of intelligence which suggests that machines could slightly surpass the human intellect, improve their own designs in ways not foreseen by their designers and, therefore, improve themselves recursively and become much more intelligent. The first improvements may be small but as machines become more intelligent, there may be a cascade of self-improvements and a sudden increase in superintelligence or singularity (Good & England, 1965).

That same year, Minsky wrote: "Once we have devised programs with a genuine capacity for self-improvement a rapid evolutionary process will begin. As the machine improves both itself and its model of itself, we shall begin to see all the phenomena associated with the terms 'consciousness', 'intuition' and 'intelligence' itself. It is hard to say how close we are to this threshold, but once it is crossed the world will not be the same" (Minsky, 1966).

Likewise, during that same year 1966, John von Neumann wrote: "There is thus this completely decisive property of complexion, that there exists a critical size below which the process of synthesis is degenerative, but above which the phenomenon of synthesis, if properly arranged, can become explosive, in other words, where syntheses of automata can proceed in such a manner that each automata will produce other automata which are more complex and of higher potentialities than itself" (Von Neumann, 1963).

But it wasn't until 1983 that the term became popular, thanks to Vernor Vinge. In January of that year, he wrote in the magazine Omni: "We will soon create intelligences greater than our own. When this happens, human history will have reached a kind of singularity" (Vinge, 1993a). A decade later, in another article he wrote: "Within thirty years, we will have the technological means to create superhuman intelligence. Shortly after, the human era will be ended" (Vinge, 1993b).

7.3 Von Neumann and the Technological Singularity

In September 1948, von Neumann gave a lecture at the Hixon Symposium in Pasadena, California, on brain mechanisms in behaviour, titled "The General and Logical Theory of Automata" (Von Neumann, 1951), whose ideas continue to be relevant today. This is due to the fact that Neumann spoke in general and abstract terms which are responsible for their timeless nature. Von Neumann's automata are a conceptual generalisation of digital computers of whose revolutionary implications he was more aware than anyone else. For him, an automaton is any piece of machinery whose behaviour may be defined precisely in strict mathematical terms. Von

Neumann's interest lay in establishing the basics for a theory of the design and functioning of these machines, which would be applicable to much more complex and sophisticated machines that any that had been built. He was convinced that this theory could be learned and understood, not only to build more capable machines but also to better understand the design and functioning of living organisms. He did not live long enough to see his insight materialise, but he did witness his theory working in living organisms. The main topic of his speech was an abstract analysis of the structure of an automaton that is sufficiently complex to be able to reproduce itself, establishing the principles of how to program a computer to replicate itself. Evidently it would make no sense to order the machine to simply reproduce as the computer's only response to this command could be: "I cannot do so, since I do not know what I am". That is to say, this is as absurd a proposition as if a woman were given a set of bottles and matrices and told to make a child. Well, according to Von Neumann, anyone who sets out to create a dynasty of machines must perform the following three actions:

1. Give the machine a detailed and complete description of itself.
2. Give the machine a second description of itself, but this second description is of a machine that has already received the first description.
3. Order the machine to build another machine that fits the second description precisely, and order the first machine to copy and pass this final order to the second machine.

For this, he showed that self-reproducing automata must have four separate components with the following functions: A component A that is an automatic factory, that is to say, an automaton that collects raw materials and processes them into an output specified by written instructions that must be provided from outside. Component B is a duplicator, that is to say, an automaton that takes a written instruction and copies it. The component C is a controller or an automaton that is hooked up to A and B. When C is given an instruction, it first passes it to B for duplication and then passes it on to A for the action, and additionally, it provides the copied instruction to the output of A while it retains the original for itself. Finally, component D writes instructions that contain a full specification causing A to build the system combining A, B and C.

Von Neumann's analysis demonstrated that a structure of this type was logically necessary and sufficient for a self-replicating automaton and guessed that it must also exist in living organisms. This was corroborated many years later by Watson and Crick (Watson, 1953) with the considerable if unacknowledged and unsolicited aid of Rosalind Franklin. Effectively, D constitutes RNA and DNA genetic material, A is the ribosomes, B is RNA and DNA polymerase enzymes and finally C is the repressor and depressor control of molecules and other items. As far as we know, the basic design of each microorganism larger than a virus is precisely how von Neumann deemed it to be.

Von Neumann's first major conclusion was that it was possible to build self-replicating automata with these characteristics. His second major conclusion, derived from Turing's work, is less well-known and makes a more in-depth study into the heart of the problem of automation. Is only, that there is a universal automaton in theory, that is a machine of a certain size and defined complications thatm given correct written instructions, is capable of performing all that any other machine can perform. From this perspective, it is not necessary to make a bigger or more complicated machine in order to accomplish more complicated tasks. All that is required is to provide longer and more detailed instructions. This universal automaton is also capable of self-replication, including it within the factory unit or element A in the aforementioned system. Von Neumann was convinced that an automaton was responsible for the possibility of an indefinitely continuing biological evolution. When evolving from the simplest to the most complex organisms, there is no need to redesign the basic biochemical machinery as you go along. All that needs to be done is to modify and extend the genetic instructions. All that is known of evolution since 1948 tends to confirm that von Neumann was right.

To conclude, it must be highlighted that according to Ulam, his namesake Stanislaw Mazur proposed the possible existence of automata that were capable of reproducing themselves if they were provided with the necessary materials, during a conversation in a cafeteria in 1929 or 1930. As may be seen, this idea is a precursor to that of von Neumann.

7.4 State of Play and Formal Aspects

For a technological singularity to take place, it is indeed essential that if an AGI is reached, that it is kept secret and therefore always under human control. However, reality shows that this is impossible. Firstly, because as the Bible says "for there is nothing hidden that will not be disclosed", one of the basic principles of cryptology today that the authors phrased in verse as followed:

> Secret of one, safe secret
> Secret of two, God keep it
> Secret of three, no secret is
> Secret of four, even the cat knows.

But perhaps the most noteworthy example of this impossibility is the legend of the Pythagorean philosopher, Hippasus of Metapontum: One of the basic and essential principles of the Pythagoreans was that the essence of all things, both in geometry as well as in the theoretical and practical affairs of human beings, may be explained in terms of "arithmos", or the properties of natural numbers and their ratios. Nevertheless, the Platonic dialogues, Theaetetus, Meno and Laws (Plato, 2002) provide convincing evidence as to how the Hellenic mathematical community was unpleasantly surprised by a discovery that, to all intents and purposes, demolished the foundations of the Pythagorean faith in whole numbers. This discovery was that some ratios, that is to say, fractions between two natural numbers such as the ratio of the hypotenuse of an isosceles triangle to one of its sides or, which is the same, to the diagonal of a square on its side, cannot be expressed in whole numbers. From there, they termed "commensurable" ratios those that could be expressed in whole numbers, which meant that both quantities had been measured by a common unit, and those that could not be expressed in this way, they termed "incommensurable" ratios. These incommensurable ratios of numbers were named "alogos" or "inexpressible", literally "alogical", although they were also called "arrestos", etymologically, that which does not make sense, "irrational". However Plato condemned the omission of this distinction between commensurable and incommensurable magnitudes in the education of young Greeks, calling it a "shameful and ridiculous ignorance",

It is not known for certain when, how and who made this discovery, although much has been written supporting the different hypotheses in this regard. A footnote in Book X of Euclid's Elements as Proposition X. 117 attributed to Proclus claims that it was discovered by the late Pythagoreans in mid-fifth century BCE. More specifically, it is attributed to Hippasus of Metapontum in the first part of the third-quarter of the fifth century. Now what we do know is that this discovery unleashed such an unprecedented crisis that the legend that followed it, although untrue, might well have been true. In the words of von Fritz:

> "It is known that the first to reveal the theory of irrational numbers to the public would perish in a shipwreck and that because the inexpressible and unimaginable should have remained hidden. Consequently, the culprit who fortuitously touched and revealed this aspect of living things was transported to his place of origin where he is whipped forever by the waves. Not a shred of pity nor an atom of compassion for one who commits a crime of lese geometry, sinning against that which is the most sacred. By revealing the secret of that which is inexpressible, he has deserved the most terrible divine punishment, and to be transported to his place of origin, that is, to nothing from whence he came, to so be divested of his being" (Von Fritz, 1945).

All of which goes to show that, if achieved, an AGI would be impossible to keep under wraps.

This may be countered by the statement that in addition to visibility, there must be the will to push the button to start the AGI. A survey conducted on this topic with a sample of more than 500 persons from all professions, nationalities and beliefs, revealed that more than 63% of men and 34% of women, the rest undefined, would push the button without hesitation.

7.5 Proposed Solution

More specifically, Polya reasoned that the discovery process in mathematics was guided by mechanisms of non-deductive inference, which consists largely of guesswork, and although they are not always accurate, they are signs of progress in problem-solving. The principles that govern this type of reasoning were called: Patterns of plausible inference. They include the following patterns:

A. Consequence verification. This type of pattern makes the consequence more credible. For example, the conjecture: "it rained last night" becomes more credible when the consequence "the street is wet" is verified.
B. Successive verification of various consequences. In this case, the modification of a new consequence counts for more or less if the new consequence differs more or less from the first verified consequence(s). For example, if we try to corroborate the conjecture: "all crows are black" and "n" crows are observed in Spain, all

of them black, then the credibility of this conjecture increases substantially if crow n + 1 is a black crow found in New Zealand, our Antipodes.

C. Verification of improbable consequences. The verification of a consequence counts more or less depending on whether the consequence is more or less probable in itself. For example, the conjecture "it rained last night" is more credible with the evidence that the "roof is leaking" than the observation "the street is wet".
D. Inference by analogy. A consequence is more credible when an analogous conjecture turns out to be true. For example, the conjecture: "of all objects with the same volume, the sphere has the smallest surface" becomes more credible when proving the related theorem that: "of all curves that have the same area, the circle has the smallest perimeter". These patterns presents modes of plausible reasoning that display conditions for making a conjecture more or less credible. This credibility depends largely on the verification of the consequence(s) associated with Hypothesis A, B and C or on the proving of analogous conjectures. Polya called this type of conclusion heuristic conclusions, which are not the true conclusions, but they serve to highlight and guide the path to those conjectures that are subsequently demonstrated on a formal basis. Polya insisted again and again that plausible reasoning is guided by rules, although these are not like the rules that characterise demonstrative reasoning, where the conclusions that are reached are always the true ones.

To compare both types of reasoning, let us first consider the "Modus Tollendo Tollens", also known as hypothetical syllogism:

$$\begin{array}{l} A \longrightarrow B \\ -B \\ \hline -A \end{array}$$

For example, "if it rains, the streets become wet", the street is not wet, ergo it has not rained.

This pattern of reasoning displays all the characteristics of demonstrative reasoning, that is to say: It is:

A. Impersonal, given that its validity does not depend on the personality, humour, tastes and so on, of the one who reasons.
B. Universal, because its form and validity are not limited to a specific domain of knowledge. It may belong to mathematics, physics, chemistry, biology, astronomy, philosophy, law, medicine and so on.
C. Self-sufficient, given that the validity of this syllogism does not depend on external aspects. Thus, once the premises are accepted, the conclusion must also be accepted.
D. Definitive, in that once the premises are accepted, one may forget them and only retain the conclusion.

Let us now compare the previous syllogism to the following heuristic syllogism:

If A——>B
B
____
A is more credible

For example, if it rains, the street becomes wet. The street is wet, ergo it has rained. The reason why the conclusion that it has rained, A, is not inferred with complete certainty is because there may be other reasons why the street is wet, B, such as the municipal cleaning services may have passed by and so on. In other terms, from A——>B and A is not necessarily deduced from B, it only makes the former more credible.

Both the demonstrative and the plausible patterns have a demonstrated and considerable similarity. Both share the first premise A——>B. With regard to the second: B false, B true are opposites with regard to truth but they are equally clear. The fundamental difference is in the conclusion(s): A false, A more credible, given that these conclusions have a very different logical status. While the conclusion of the demonstrative pattern is at the same level as its premises, that of plausible reasoning is of a different nature, less precise and open to discussion.

That said, it may seem that heuristic syllogism does not fulfil the characteristics of the demonstrative pattern that we have just pointed out: Universality, Self-sufficiency and Definitiveness. Nevertheless, for Polya, to a certain extent, these characteristics are present in the pattern of plausible reasoning. Effectively, for all rational persons, the heuristic conclusion A is more credible based on the premises than without them. In this sense, the rule: "Consequence verification strengthens the conjecture" is impersonal and universal as its validity only depends on its form and not its particular content. Nevertheless, when going more into detail and enquiring how much more credible is the conclusion or what is the weight of the evidence, personal differences and different domains of knowledge enter into play. Differences of prior knowledge, of experience in a particular topic, of style or personal opinion, are all determining factors when judging the degree of credibility of a conclusion as well as the weight of the evidence. With regard to self-sufficiency and definitiveness, the heuristic conclusion is based on the premises and does not depend on external aspects. Nevertheless, if there is new information in the future, for example, during a period of drought, the credibility of the conclusion may change until the point that it is false; for example, it may not have rained. Thus, it may be said that based exclusively on the information of the premises, the conclusion is justified. Only that this conclusion is provisional, that is, it is not lasting or definitive. The verdict of a jury may convict an innocent individual or exonerate a criminal, but the injustice of the verdict is justified, perhaps only logically, in the sense that based on the evidence of the available information, no other verdict was possible.

To summarise, plausible reasoning, without achieving the force of a strict logical demonstration, is necessary in scientific conclusions and in daily life, when passing from observations and experimental and/or empirical data to unobserved values.

The main interest then lies in how to pass from logical, deterministic and certain reasoning, to plausible, uncertain reasoning, given that this reasoning reflects the current status of a class of knowledge that goes from pure common sense to sophisticated language currently used to build scientific and human knowledge in all their impressive variety of forms. Among these forms are those used in philosophy, medicine, science or law, among others. This makes it necessary, convenient and urgent, given that uncertainty is a ubiquitous phenomenon, to change the form of reasoning to include plausible, conjectural or hypothetical reasoning that is rigorously treated.

For this, let us consider the following two syllogistic rules:

(A) "Modus Ponens": If A is true, then B is true
A is true

Ergo, B is true

(B) "Modus Tollens": If A is true, then B is true
B is false

Ergo, A is false

Certainly, there are many other rules of demonstrative inference, such as disjunctive syllogism, addition and so on, but they may all be reduced to earlier syllogisms which prevents considering them here and now.

Each one of the strong syllogisms or earlier demonstratives, (A) and (B), give rise to weak or plausible syllogisms or the following heuristics:

(A′). If A is true, then B is more plausible or credible.
B is true.

Ergo, A is more plausible or credible.

For example, A may be: There was a hailstorm at 6 in the evening and B the sky was cloudy and grey at 5:30 in the evening.

(B′) If A is true, then B is true.
A is false.

Ergo, B is less plausible or credible.

As is the case in the previous example.

Therefore, Polya's essential contribution was to emphasise the qualitative aspects of the extension of classic Aristotelian schemas, by proposing new schemas that serve to support plausible reasoning. As a matter of fact, Polya (1954) provided a set of plausible inference patterns that are used in common sense reasoning, police investigations, in experimental and observational sciences such as astronomy, law, medicine and so on, used to discover new truths. From Polya's qualitative aspects, based on Cox's axioms (Cox, 1946) which let us represent the degrees of belief that satisfy the rules regarding the combination of probabilities, it is possible to quantitatively model plausible reasoning. For this, essentially, each proposition to be reasoned is assigned a degree of credibility and when new evidence is received, this assignment must be modified to take this new information into account.

Cox's hypotheses are the following:

- The degrees of plausibility are represented by means of real numbers. Symbolically, each proposition A/B is associated with a number p(A/B), which will measure its degree of plausibility.
- The degree of plausibility of each inference, given certain evidence. Symbolically: p(notA/B)g(p(A/B)).
- The degree of plausibility of the conjunction of two inferences, given certain evidence, is determined according to the degree of plausibility of the other inference, given the evidence and if the other inference is true. Symbolically: p(AB/C)=f(p(A/C),p(B/AC)).

Based on these axioms, it has been demonstrated that non-negative numbers are assigned to sentences A/B that satisfy the basic rules of calculating the following probabilities:

1. Product rule: p(AB/C)=p(A/C).p(B/AC)=p(B/C).p(A/BC).
2. Sum rule: p(A/B)+p(A/B)=1.

Once we obtain the representation of the degrees of plausibility by means of conditioned probabilities, we may implement deductive logical reasoning and the rules of plausible reasoning, thanks to Polya. Effectively, if C is designated as the proposition "If A then B", we see that the Modus Ponens and Modus Tollens are adapted respectively, to the product rules:

$$p(B/AC) = p(AB/C)/p(A/C) \text{and} \, p(A/notBC) \\ = p(AnotB/C)/p(notB/C).$$

From the Modus Ponens and Modus Tollens we have respectively, p(AB/C)=p(A/C) and p(AnotB/C)=0, therefore p(B/AC)=1 and p(A/notBC)=0. Therefore, the Modus Ponens and Modus Tollens may be modelled with the help of probabilities as extreme cases where we are sure of the conclusions. Even more, given that is possible to have quantitative versions of the weak or plausible rules. Thus for example, the previous weak syllogism:

If A is true, then B is true.
B is true.
———————————————————.
Ergo, A is more plausible

corresponding to the product rule in the form: p(A/BC)=p(A/C).p(B/AC)/p(B/C). Now, from Modus Ponens p(B/AC)=1, and as p(B/C)< o = 1, we have p(A/BC)> o =p(A/C), as indicated in the plausible syllogism. In turn, the syllogism:

> If A is true, then B is true.
> A is false.
> ————————————————————.
> Ergo, B is less plausible,

corresponding to the product rule in the form: p(B/notAC)=p(B/C).p(notA/BC)/p(notA/C). From the previous conclusion, it follows that: p(notA/BC)< o =p(notA/C); therefore p(B/notAC)< o = p(B/C), as indicated in the syllogism.

Finally, in relation to the plausible syllogism:

> If A is true, then B is more plausible.
> B is true.
> ——————————————————————————.
> Ergo, A is more plausible.

If C designates the available information, the first premise is p(B/AC)> o = p(B/C) that applied to the product rule: p(A/BC)=p(A/C).p(B/AC)/p(B/C), directly gives: p(A/BC)> o =p(A/C).

As may be seen, it is possible to model logical reasoning and patterns of plausible reasoning in probabilistic terms, and therefore we may effectively speak of probabilities as extended logic.

To sum up, Cox's axioms let us introduce probability as a measure of the degrees of plausibility or belief regarding a proposition based on the information available at a given moment. In the presence of additional information, the probabilities are updated according to Bayes' formula.

7.6 Development and Conclusion

In 1763, a work by the Reverend Thomas Bayes was published posthumously, having been sent by his friend and colleague Richard Price to the journal of the Royal Society of London. In his preface to the submission, Price wrote: "I now send you an essay which I have found among the papers of a deceased friend Mr Bayes, and which, in my opinion, has great merit, and well deserves to be preserved. Experimental philosophy you will find is nearly interested in the subject of it, and on this account there seems to be particular reason for thinking that a communication of it to the Royal Society cannot be improper."

> […] In an introduction which he has written to this Essay, he says that his design at first in thinking on the subject of his was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumstances, upon supposition that we know nothing concerning it but that, under the same circumstances.

Apart from the introduction, the work has two sections. In the first section, Bayes presents his theory, defining probability as a subjective relationship between a certain present and an uncertain future, for example, between a certain sum to be paid today in order to receive a random sum were a given event to take place. From there, Bayes makes a rigorous demonstration of the theorems of total probability and of a compound probability. The second section includes the demonstration on the inversion of probability, which constitutes the solution to the problem originally posed.

Bayes did not extend his results beyond uniform distribution. This would be done later by Pierre Simon de Laplace, but his vision of probability and abductive inference has been widely adopted and applied to large number of problems in statistical inference and in decision theory and so on. And this is because this theorem provides a response to the following important question: How can a person update their current belief when they discover new evidence, for example, from an experiment?

Bayes' work remained ignored for a decade on the shelves of the Royal Society and had little or no influence on mathematicians of that period, until it was rediscovered by Jean-Antoine Nicolás Caritat, Marquis of (Condorcet, 1970), and its consideration by Laplace who was responsible for its most well-known form.

Since then and long before the advent of computer science, Bayes' theorem has had a critical influence in scientific development, as it has made a significant contribution to the assessment of the degrees of certainty in different hypotheses. It has had an enormous effect and multiple and varied applications in the most diverse fields and domains such as law, medicine, science, IT, games, control systems and genetics. For example, and as a case study, it is the mechanism followed in genetic identification tests which conduct measurements based on the presence or absence of certain alleles to make an inference, thus obtaining a certain numeric value which measures the probability that two samples belong to persons who are related by blood, and which, above a certain value, is even considered as conclusive proof in a legal process.

Thus, even beliefs may be measured such that they may become certainties for all intents and purposes. This leads to the consideration that Bayes' theorem provides a quantitative measurement to improve the principle of abduction which may adopt the following form: "If we observe a particular case or evidence x consistent with the conjecture of hypothesis h, then the probability that h is true increases by a certain value".

Although the formula, equation or rule that is obtained from Bayes' theorem is relatively simple, its results are not at all intuitive because it plays with conditional probabilities. As a matter of fact, Bayes' theorem is not easy to understand or learn and apply as often its results are not believed since being so less intuitive, they are simply assumed to be errors or incomprehensible statistical gibberish. One example will demonstrate this counterintuitive nature of Bayes' theorem: The Garcias are a family with two babies. One of them is a girl. What is the probability that the other is also a girl? The most frequent response is 1/2, but it is incorrect. The mental model theory explains this error on the grounds that people interpret the problem as asking for the probability of a baby being a girl and consequently build a girl, boy model, from which they deduce that the probability of it being a girl is half. However, what the problem really demands is a conditional probability: P(both are girls/one is a girl). A simple mathematical justification of the solution would be the following: The sample space is expressed as follows: = {(boy, boy), (boy, girl), (girl, boy), (girl, girl)}. Now, if we know that one is a girl, the space would be = {(boy, girl), (girl, boy), (girl, girl)}. Therefore, admitting that both sexes are equally probable, we would have: P (two girls/at least one is a girl) = 1/3. As we can see, it is easy to commit errors when dealing with conditional probability owing to the cognitive functioning of human mental patterns.

The three main characteristics that differentiate Bayesian methods from other methods and techniques of statistical analysis are the following:

One, the analyst must quantify their judgements in numerical terms. From this point of view, Bayesian analysis helps to quantify what Polya termed plausible reasoning, which emerges from forcing arguments of classical logic to give rise to plausible heuristic patterns. It has a close connection with abductive logic. Two, the analyst does not take the available information as a given, and on this basis develops their conclusions on the relative merits of other hypotheses; rather they investigate in stages the truth of each working hypothesis, focusing only on the plausibility of each new element of evidence for each working hypothesis if these were true. Three, the judgements are constructed on the basis of pieces of evidence. It is unlike the summary of evidence to be made when judging their significance for the final conclusions. Probability calculation summarises by saying: If this is the reading of the evidence at hand, this is the conclusion that follows from them.

Frequentist statistics is based on the idea of quantifying the probability of an event from the relative frequency of its appearance. Bayesian statistics, on the other hand, is based on the notion that probability represents the degree of belief in the event in question. The difference between the two is therefore a question of the concept of probability. For classical or frequentist statistics, probability is an objective concept which is found in nature, whereas Bayesian statistics is found in the observer, and is thus a subjective concept. Thus, classical statistics takes only the samples obtained as the source of information, whereas in Bayesian statistics, apart from the sample, prior or external information that is available in relation to the phenomena to be modelled also plays an essential role.

The following important concepts of Bayes' theory are of interest here: One, inverse probability, which states that the probability of a cause given an effect is different from the probability of an effect given a cause. For example, let us assume that A denotes the event or case that a person is a woman, and B the event that she is pregnant. It may be observed intuitively that: P(A/B) or the probability that a person is a woman given that she is pregnant is 100%, given that all pregnant persons are women. Nevertheless, the P(B/A) or probability of being pregnant given that she is a woman is only 2%. Bayes found a connection between both probabilities, which is highly important because by knowing only one of the two probabilities, we may know the other.

The probability calculation is therefore: $P(A/B) = P(A\&B)/P(B)$. That is to say, the probability of an event A, given another event B, is equal to the probability of the intersection of both events divided by the probability of event B. By subtracting from the previous expression the probability of the intersection, we have: $P(A\&B) = P(A/B)\times P(A)$ [1]. Similarly, we have: $P(B/A) = P(B\&A)/P(A)$ and therefore, $P(B\&A) = P(B/A)\times P(A)$ [2]. As the probability of the intersection is commutative, $P(A\&B) = P(B\&A)$, by equating the expressions [1] and [2], we obtain: $P(A/B)\times P(B) = P(B/A)\times P(A)$ and subtracting from this expression P(B/A), we obtain:

$P(B/A) = P(A/B) \times P(B)/P(A)$, which is no more or no less than Bayes' formula. In this formula, P(A) is total probability, P(B) a priori probability, P(A/B) is the conditional probability of an event A conditioned to another given event B, and finally P(B/A) is the a posteriori probability.

Two, the theory of plausibility. From the previous step, we are now able to undertake the complexities of Bayesian analysis; that is to say, that which deals with the relationship between probability and plausibility. This relationship, now expressed in terms of hypothesis or conjecture and evidence, is expressed by the three following equivalent equations:

$$P(h/e) = P(e/h) \times P(h)/P(e) = P(e/h) \times P(h)/$$
$$\left[P(e/h) \times P(h) + P(e/h') \times P(h')\right] = LP(h)/LP(h) + P(h')\left.\right].$$

where:

P(h) is the a priori probability which represents someone's uncertain knowledge regarding the occurrence of a given hypothesis or conjecture h. This probability is obtained from all the information available in this regard.

P(e/h) is known as plausibility or belief in e given h. As h is a fixed value and e takes values within its domain, P(e/h) is not interpreted as a conditional probability. Therefore, this function is denoted by L. This plausibility function.

L(h/e)=P(e/h) must be deemed a function of h for a fixed value of e, but additionally, P(e/h) represents the researcher's degree of belief that the data, evidence, take the value e given the hypothetical information that the hypotheses take certain h values. The values (h/e)=P(e/h) should be obtained from the conditional probabilities of e for the different values of h. Therefore, this function must be such that an h for which P(e/h) takes a high value denotes that on one hand, this h has a high probability of being true compared to another case h′ for which the P(e/h′) is lower; and on the other hand, that the occurrence of e is more plausible when the P(e/h) value is higher.

That is to say, in terms of hypothesis/evidence, a plausibility function must establish that the hypotheses h that have a greater probability without plausibility due to e must somehow have a higher probability when the occurrence of h is observed. Let us translate all of the above into a concrete example. If a person looks out of the window of their house and sees a metal artefact with four wheels, why do they think it is a car? Because these elements appear in all cars. There may be some doubt regarding whether it is a motorbike, but surely nobody thinks it is a donkey. Formally, if the knowledge were modelled as follows:

- e is the data establishing the presence of metal and four wheels.
- h is a variable, hypothesis or conjecture, with values:
  V(h) = {h1=car, h2=motorbike, h3=donkey}.

Then the following plausibilities may be established:

- The statement “cars generally look like this” is probably quite true, which implies that P(e/h1)#1.
- The sentence “Donkeys generally look like this” implies that P(e/h3)#0.
- The assertion “motorbikes generally look like this” implies that P(e/h2) is closer to P(e/h1) than P(e/h3).

It is worth pointing out that P(e/h1) provides two items of information. Firstly, that cars are the most credible hypothesis. Secondly, that the observed data are most plausible if it was a car that was observed.

Three, the normalising constant. Retaking the equation: P(h/e)=P(e/h)×P(h)/P(e), pending comment the term P(e). This value is a constant as it is a fixed value in the expression. There are two ways by which this value may be obtained:

- Considering the joint probability P(h,e) and calculating the marginal value of P(e).
- Or applying the theorem of total probability which establishes that, given the two variables h and e:

$$P(e) = \text{Sum over h of } P(e/h) \times P(h),$$

and replacing it in Bayes’ rule, we have:

$$P(h/e) = P(e/h) \times P(h) / \text{Sum over h of } P(e/h) \times P(h).$$

This equation is highly interesting given that an a posteriori probability may be calculated using only an a priori probability and a plausibility function. That is to say, once all the numerators in Bayes' rule to calculate P(h/e) are known, for all possible values of h, only the variable h has to be eliminated by sums to obtain the denominator.

As P(e) is a constant, Bayes' rule may also be expressed as follows:

> P(h/e)=@ P(e/h)×P(h), where @ is the normalising constant that is calculated by the theorem of total probability.

A specific case of Bayes' rule is when the a priori probability of the different hypotheses, conjectures, causes or scenarios P(h), are the same, which leads to:

> P(h/e)=kP(e/h), with k a constant that is independent of each h. That is to say, if there are no reasons to prefer some hypotheses over others, their probabilities are proportional to their plausibility.

Taking into account the basic characteristics of intelligence given in its definition of the chapter 1: creativity, disambiguation, enantiodromy, holism, inferences and reasoning, learning, proactivity, troubleshooting, and teleology, and the differents evidences, programs and systems in AI that work correctly: Alfazero, AlfaGo, Deep Blue, Deep Mind, EQP, IAMUS, Jeopardy, Libratus, Mitsuko, SAM, and differents traductors, following to Lara and colleagues (Lara et al. 2019), the following can be concluded:

– With a probability of the 85% chance, the AIG will come true. However, this does not imply that the date on wich it will occur be established by a simple rule of three.
– Certainly, so far, virtually all the achievements in AI, were based on digital computers, notwithstanding, nothing indicating, but quite the contrary, that the AIG as well will be supported exclusively by them.
– To make a pertinent analogy, the AIG is currently at the same point as the digital computer before Turing proposed its TMU. There were many particular TMs that solved specific problems, but there was a lack of another TM to integrated them. Said colloquially, all, or almost all, of the wickerwork are available, but the final touch that allows the basquet to be made is missing.

7.7 Ethical Issues

Plato, in his dialogue titled *Euthydemus*, states: "I would like to know which is the defining characteristic of piety which makes an action to be considered as pious (…) in order to observe it and use it as a standard to judge your actions and that of others". This statement, affirmed by the character of Socrates in the Plato dialogue, evidences the search for an algorithm that allows to tell piety from impiety. In his *Nicomachean Ethics*, Aristotle went a step further and re-stated the idea of his master in more precise terms:

> "And we deliberate not about ends, but about means. A doctor does not deliberate whether he is to cure his patient, nor an orator whether he is to convince his audience, nor a statesman whether he is to secure good government, nor does anyone else debate about the end of his profession or calling; they take some end for granted, and consider how and by what means it can be achieved (…) until they reach the first link in the chain of causes, which is the last in the order of discovery. (…) Then, if they have come up against an impossibility, they abandon the project—for instance, if it requires money and money cannot be procured; but if on the other hand it proves to be something possible, they begin to act." (Aristóteles, n.d.).

Considering the above, this text addresses the ethics and the ethical issues entailed by the development of AGI. For the purposes of this text, ethics is defined as the set of principles that govern one's intended behaviour. However, considering that the authors are not philosophers and do not wish to bite off more than they can chew, this issue shall be considered from a scientific and pragmatic point of view as used in software and artificial intelligence (AI) development. That is, in practical terms: Are we doing the right thing? The things that are being done, are being done the right way?

In the first place, the difference between ethics and metaethics is to be established. The subject of ethics is what is right and what is wrong, and which moral guidelines should govern one's behaviour and that of others. A statement such as "killing kittens is wrong" belongs to the realms of ethics. However, metaethics takes one step backwards and wonders: What does it mean to affirm that something is "right" or "wrong"? Why one set of principles has to be adopted instead of another wholly different set? "The ethical system adopted must be based on improving the wellbeing of sentient beings" is a metaethical statement from which "killing kittens is wrong" can be deduced.

Hume, in its *A Treatise of Human Nature, (Hume, 1739-1740)*, complained that most philosophers had a penchant for suddenly stating **what should be true** when previously they had limited to describe **what was true**. What is more, considering **what it should be** is something completely different from considering **what it is.** The first consideration entails a judgement, as it states what should be the case, while the second consideration is purely descriptive, and affirms what, in fact, is. With respect to this, there is a proclamation that states: **What it should be cannot be derived from what it is.**

This poses a scientific challenge, considering that **what it should be** cannot be derived from **what it is,** we have a problem, since **what it is** is all we have. That is, there is nothing outside the natural world to whom to turn in search of orientation about how to behave.

All efforts to revert this situation are futile for the simple reason that nature does not judge, nor provides guidelines, nor knows or cares for what it should be. Humans, although part of nature themselves, have been granted the capacity to judge, but different persons may reach divergent and even opposing conclusions.

In order to understand why **what it should be** cannot be derived from **what it is**, it is useful to consider how to derive anything from anything different. For these purpose, the best tool, in logical terms, is to be used: Syllogism, and which can be expressed in general and particular terms as follows:

1. X is true. All men are mortal.
2. If X is true, then Y is true. Dan is a man.
3. Therefore, Y is true. Dan is mortal.

The two first statements of a syllogism are its propositions, and the third is its conclusion. In logics, it is stated that an argument is valid when its conclusion can be logically deduced from its propositions. In contrast, it is said that an argument is sound when its conclusion can be logically deduced from its propositions *and* those are true; this is a much more restrictive condition.

Well, if **what it should be** was attempted to be derived from **what it is** using a syllogism, something similar to the following argument may result.

1. I'd like to have the last slice of pie.
2. If I don't hurry up, someone will come and take the last slice of pie.
3. Therefore, I must hurry.

While the above argument may seem a good argument, it is not a logically valid syllogism. The propositions are two statements describing **what it is**: My wish to have the last slice of pie and the possibility of losing my chance if someone snatches it before me are both objective statements about the world, regardless of whether they are true or not. And the conclusion is, undeniably, a statement of **what it should be (done).** But a closer look, going beyond the usual meaning of words and deeper into their underlying logical contents, reveals that something is missing. Propositions (1) and (2) do not entail conclusion (3). In fact, they entail: Therefore, if I don't hurry up, I won't be able to have the last slice of pie.

For the original conclusion to be validly derived, a new proposition should be added, something along the lines of: 2a. I should act in such a way that I get what I'd like to have.

With this addition, the argument becomes valid. And it no longer proposes to deduce **what it should be** from **what it is**, since the new proposition directly introduces a **what it should be** argument. It has only derived a proposition **what it should be** from another similar proposition plus a set of **what it is** propositions.

This is the problem of trying to derive **what it should be** from **what it is**: It is impossible. If someone states that they have derived **what it should be** from **what it is**, it is the same thing as if they state that they have added to even numbers and obtained an odd number as a result.

A modern variation on the issue of obtaining **what it is** from **what it should be** consists on affirming that morals can be reduced or submitted to scientific practice, and this is an idea that is usually presented approximately as below:

1. Achieving X would make the world a better place.
2. Science can tell us how to achieve X.
3. Therefore, we should do as science says.

In this case, the "hidden" proposition is.

2a. The world should be made a better place.

This may seem a tautology, depending on how the term is defined. But even if the hidden proposition is included, or even buried in the definition of "better", the argument still positively states that something must be done. Such statements cannot be based exclusively in objective statements. Who decides what is "better"?

Defenders of this fallacy state, sometimes, that they only pose reasonable propositions; that this is what science does all the time, and they do nothing different. However, such an argument overlooks an important aspect of what science is. Consider the following statements:

- The universe is expanding.
- Humans and chimpanzees share a common ancestor.
- Efforts should be made to ensure that people live longer, happier lives.

Although all the above statements are true, only the two first are scientific. The reason of this is that both may have been refutable, that is, that either of them could had been refuted or proved wrong. They are not true by definition or by supposition. One could imagine possible worlds in which the universe remained static, or on which there existed species similar to humans and chimpanzees had not evolved from a common ancestors. To decide whether such statements are true or not, they are empirically tested. However, no one even thinks of carrying out an essay to decide whether efforts should be made to ensure that people lived longer, happier lives. It is assumed that it is so, or it is attempted to be derived from related propositions. This vital "additional ingredient" summarises the difference between how science works and how thought about right or wrong is developed. Science does require propositions: There are certain epistemological affirmations, such as trust in basic information apprehended by senses, that have an important role in constructing sets of stable beliefs for scientists. However, propositions needed for science to function do not have the same effect in morals.

This does not mean that **what it should be** is not an issue that should be pondered using reason and rationality. There is a form of logical thinking called instrumental rationality whose purpose is to answer questions such as: If one wishes to attain certain goal, what is it to be done to attain it? Then, the question is derived to deciding which is such goal.

Lack of a fundamental objective scientific basis for morality can be disturbing. But, unfortunately, that is what it is.

Certainly, science has little to say about ethics beyond maybe a few inspiring remarks. However, it does have something to say about metaethics: that is, that human ethical systems are human constructs, not systems that were out in the world of their own, and that they should be appraised accordingly. In fact, there are two ideas that can be used as a helpful starting point: Consequentialism that states that the moral implications of an action are determined by the consequences that such action has. And deontology, derived from the Greek words for duty (*deon*), being (*ontos*), thus, together, "what it should be", and knowledge or study (*logos*).Well, deontologists believe that acts are in themselves right or wrong, regardless of the consequences that they may have. An example of the first line of thinking would be the motto of utilitarianism: "the greatest good for the greatest number". And an example of the second line of thinking would be the Golden Rule: "treat others as you wish to be treated". Both, at first consideration, seem to be excellent ideas; the problem is that they often conflict and are contradictory.

The idea that moral principles are constructed by human beings based on their own subjective judgements and beliefs, and not based on something external to them, is known as moral constructivism. A moral constructivist recognised that morals arise from individuals and societies but accept that such individuals and societies shall consider the resulting set of beliefs as "right" and shall judge others in consequence. Moral constructivists have no qualms in telling others that they are in the wrong. Besides, the fact that morals are a construction does not equate that they are arbitrary. Ethical systems have been invented by humans, but humans also can engage in productive conversations on how to improve such systems, in much the same way as with anything created by humans.

There are two types of constructivism: Kantian constructivism and Humean constructivism. Kant was a textbook deontologist and based his opinions on morality on categorical imperative: "Act only according to that maxim whereby you can, at the same time, will that it should become a universal law" (Kant, 1899). Hume, on its part, rejected absolute moral principles and stated that reason can help to attain what one wants, but that what one wants is defined by one's passions. A Kantian constructivist accepts that morals are constructed by humans, but believes that all rational persons would construct the same moral framework if they only thought sufficiently clearly on the subject. A Humean constructivist goes beyond that and states that different persons may very well construct very different moral frameworks.

Considering historical evolutions, Hume was right. Humans do not have any objective guidelines to distinguish right from wrong: Not from

God, not from nature, not from reason itself. Judging what is right and what is not is an intrinsically human action, and this has to be faced. Morality only exists in as much as it is made to exists, and others could not stand by the same principles as one does. Once established that ethical standards are a human construction, what consequences does this have for traditional AI and AGI? With respect to the first one, in a lecture held on January 2017 in Asilomar, California, and which was attended by most people who have something to say regarding AI development, a series of ethical principles for the development of AI were established:

- Security. AI systems must be secure throughout their entire operational lives and verifiable whenever this is possible and feasible.
- Failure transparency. If an AI system causes harm, it must be possible to find out why.
- Legal transparency. Any intervention of an autonomous system in decision-making must be accompanied by a satisfactory explanation, which may be revised by a competent human authority.
- Accountability. Developers of advanced AI systems are not oblivious of the moral implications of the use and abuse and actions of such systems, and have the opportunity and the responsibility of ending such implications.
- Value conformity. Highly autonomous AI systems must be designed in such a way that it may be guaranteed that their goals and behaviours are compliant to human values while they are in operation.
- Human values. AI systems must be designed and managed to ensure that they are compatible with the principles of human dignity, human rights and freedoms and cultural diversity.
- Personal privacy. Persons must be entitled to access, manage and control the data generated by them, considering the capacity of AI systems of analysing and using this information.
- Freedom and privacy. Application of AI to personal data cannot restrict unjustifiably real or perceived personal freedom.
- Shared benefits. AI technologies must benefit and empower as many people as possible.
- Shared prosperity. Economic prosperity rising from the use of AI must be widely shared for the benefits of all humankind.
- Human control. Human beings must be able to decide whether they delegate their decisions on AI systems to achieve previously chosen goals and in which way.

- No disruption. The power granted by control of highly advanced AI systems must respect and improve civil and social processes on which societal health relies, and not subvert them.
- Arms race. All types of arms races regarding lethal autonomous weapons must be avoided.
- Precaution regarding capacity. Since there is no consensus with respect to this, strong presumptions on upper limits of future AI capacities must be avoided.
- Importance. Advanced AI may represent profound change in the history of life on Earth and must be planned and managed considering the corresponding significance and resources.
- Risks. Risks associated to AI, especially catastrophic risks or existential risks, are subject to planning and mitigation efforts according to their potential impact.
- Recursive self-improvement. AI systems designed with recursive self-improvement or self-replication capabilities so that they may lead to a rapid quantitative or qualitative improvement must comply with strict security and control measures.
- Common good. Superintelligence capabilities must be only developed at the service of shared ethical ideals and in the benefit of the entire mankind, not of a single State or organisation.

Incidentally, the panel on long-term AI included figures as relevant in the AI field including Musk, Russell, Kurzweil, Hassabis, Harris, Bostrom, Chalmers, Selman, Tallinnm and (Tegmark, 2017), who organised the conference.

To finish, it must be stated that authors consider these principles to be "all pie in the sky". In fact, if an AGI was completed, it would be autonomous and smarter than humans, and it would, eventually state its own ethical principles, in almost all certainty not be the ones stated above.

References

Adams, H. (1920). The rule of phase applied to history. In, H. Adams & B. Adams. The degradation the democratic dogma. Macmillan. PP 267–311.

Aristóteles. (n.d.). *Nicomachean Ethics*. www.perseus.tufts.edu

Butler, S. (1872–R.E. 1901). Erewhon. Trubner&Co.

Cellarius. (1863, June 11). Darwin among the machines. The Press Christchurch, 180–185.

Condorcet, N. (1970). Esquisse d'un tableau historique de l'esprit humaine. Vrin.
Good, I. J., & England, C. B. (1965). *The estimation of probabilities.*
Hume, D. (1739–1740). *Treatise of Human Nature* (3 Vols). John Noon.
Kant, I. (1899). *Kant's Critique of Practical Reason and Others Works in the Theory of Ethics.* Logman Green, and Co.
Minsky, M. (1966). Artificial Intelligence. Scientific American *215*(3), 257.
Plato, R. (2002). A Parameter Estimation Problem and its Regularization by the Conjugate Gradient Method. In *Inverse Problems in Engineering Mechanics III* (pp. 245–250). Elsevier Science Ltd.
Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence*. Vintage.
Thornton, R. (1847). The age of machinery. The expounder of primitive Christianity 4, 281.
Vinge, V. (1993a, March). Technological singularity. In *VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute* (pp. 30–31).
Vinge, V. (1993b). The coming technological singularity: How to survive in the post-human era. Science fiction criticism: An anthology of essential writings, 352–363.
Von Fritz, K. (1945). The discovery of incommensurability by Hippasus of Metapontum. *Annals of mathematics*, 242–264.
Von Neumann, J. (1951). The general and logical theory of automata. Pasadena, CA, USA, September 1948. *LA Jeffress (ed.).*
Von Neumann, J. (1963). *The general and logical theory of automata* (Vol. 5).
Watson, J. D. & F. H. C. Crick (1953). A structure for deoxyribose nucleic acid. Nature 171, n° 4356, 737–738.

Conclusions

This book is aiming to be a baseline for further articles, papers, readings, professionals and researches. Further academic and professional works supported with quantitative and market research methods would be the next steps for highlighting Artificial Intelligence (AI) impacts in businesses. By its nature, most of the references cited are from the latest developments of scientific researches and from the business life but the findings and consequences are consistent with the previous literature on innovation and technological impacts on business, legal, engineering, economy and daily life.

As a summary, the current impacts and the expected disruptive changes of the AI and robotics on the economics and business which are the earlier stages of the "IV Industrial Revolution" would be in the near future as solutions particularly interesting in markets, with a large quantity of data and recurrence in service needs, like retail, healthcare services and public administrations.

AI is reshaping economies, and promises to boost productivity and improve efficiency. These technologies, however, are still in an emergency phase, but they are ready to address global challenges and promote innovation and social benefits.

For that reason, although there are different speeds in adopting AI solutions by sectors, however there is an enormous opportunity in adopting AI. The use of AI in all sectors will grow exponentially in upcoming years, creating solutions that we cannot even imagine still.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9

Glossary of AI for Business

Big Data: data that contains greater variety, arriving in increasing volumes and with more velocity.

AI ethical issues: for the purposes of this book, ethics is defined as the set of principles that govern one's intended behaviour. However, considering that the authors are not philosophers and do not wish to bite off more than they can chew, this issue shall be considered from a scientific and pragmatic point of view as used in software and AI development.

AI readiness framework: a framework intended to assist efforts to address one of the first of these challenges: assessment of organisational AI readiness, that is an organisation's ability to deploy AI technologies to enable digital transformation, in four key dimensions: technologies, activities, boundaries and goals.

Algorithms: advanced computer programs, which, in companies, by gathering workers' data make it possible to manage an intrinsic and increasingly important part of the labour relationship by means of automated decision-making.

Approaches to working with AI currently: the symbolic approach, Neural networks and the Bayesian approach.

Artificial Intelligence (AI): a loose term used to describe a range of advanced technologies that exhibit human-like intelligence including machine learning, autonomous robotics and vehicles, computer vision, language processing, virtual agents and neural networks.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9

Autonomous Vehicle (AV): a vehicle capable of sensing its environment and operating without human involvement using various in-vehicle technologies and sensors.

Business benefits of AI: Improving personalised shopping experiences, automating customer interactions, real-time assistance, data mining, predicting outcomes and improving the recruitment process.

Chatbot: a computer program designed to simulate conversation with human users, over the internet, which boosts operational efficiency and brings cost savings to businesses while offering convenience and added services to internal employees and external customers.

Crowdworking: a new generation of spaces, open to public and private partners, that will become the centre that brings together the local ecosystem.

Deep learning: a type of machine learning and Artificial Intelligence that imitates the way humans gain certain types of knowledge.

Digital business strategy: organisational strategy formulated and executed by leveraging digital resources to create differential value.

Digital disconnection: this regulation puts a limit on the use of business communication technological media in rest periods and it demands the respect of the maximum hours of the working day.

Digital Framework: a framework that allows the company to addresses digital goals and objectives. To cast an overall vision for what digital will be doing for the company.

Digital Scope: to address the company's approach to key areas of digital and outline the purpose, key initiatives and challenges of each objective.

Digital Strategy Plan: the articulation of an organisation's vision, goals and purpose for engaging with digital solutions and technologies. It articulates the opportunities and challenges related to digital activities, the governance and management arrangements and risk management issues. It lays out a plan of action in order to maximise the business benefits of digital initiatives to the organisation.

Digital Transformation Journey: a phase-based focus with a clear roadmap that involves a variety of interested parties, beyond the silos and limitations, internal and external.

Digital Transformation Strategic Framework Elements: Strategy, technology & Data, Client Experience (360°) and Operational Processes.

Digital Transformation Strategy Objective: to create the capacity necessary for taking maximum advantage of the possibilities and opportunities offered by new technologies and their impact in a more rapid, better and more innovative way in the future.

Digital Transformation Strategy: the process of identifying, articulating and executing digital opportunities that extend the competitive advantage of the organisation.

Ethical principles for the development of AI: Security, Failure and Legal transparency, Accountability, Value conformity, Human values, Personal privacy, Freedom and privacy, Shared benefits and prosperity, Human control, No disruption, Arms race, Precaution regarding capacity, Importance, Risks, Recursive self-improvement and Common good.

Human AI characteristics: Awareness and recognition of our surroundings, Interpreting and ascribing "meaning" to information, Handling ambiguity, Applying judgement and Making decisions.

Industrialisation of AI: Expert Systems and the Rehabilitation of Neural Networks (1969–1990).

Machine Learning design in business: a catalyst instrument for enhancing business scalability and improving business operations for companies across the globe. Businesses using approaches such as creating a "workbench" for data science innovation or providing a "governed path to production" which enables "data stream model consumption".

Machine Learning: a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. It is used in internet search engines, email filters to sort out spam, websites to make personalised recommendations, banking software to detect unusual transactions and lots of apps on our phones such as voice recognition.

Neural Networks: series of algorithms that endeavours to recognise underlying relationships in a set of data through a process that mimics the way the human brain operates. By adopting Artificial Neural Networks businesses are able to optimise their marketing strategy.

Ride Sharing Company: a transportation network company (ride-hailing service) that, via websites and mobile apps, matches passengers with drivers of vehicles for hire.

Robotic Process Automation (RPA): a software technology that makes it easy to build, deploy and manage software robots that emulate humans actions. A revolution in business process automation.

Space Economics: the full range of activities and the use of resources that create value and benefits to human beings in the course of exploring, researching, understanding, managing and utilising space.

Technological Singularity: the possibility of achieving an artificial general intelligence that surpasses human intelligence, one of the major paradigms of life for humanity today.

Teleworking: a form of organisation of work, which responds to both technological innovation—the new information and communication technologies—and organisational flexibility in a changing environment.

Work-related Algorithms: staff recruitment, the organisation of work schedules, the monitoring and tracking of activity, professional promotion, the calculation of performance, the application of the disciplinary system, the determination of salaries or even the evaluation of a dismissal.

Index

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

A. Landeta Echeberria (ed.), *Artificial Intelligence for Business*,
https://doi.org/10.1007/978-3-030-88241-9