



HAL
open science

Introduction

Andrew King, Nicolas Duchateau

► **To cite this version:**

Andrew King, Nicolas Duchateau. Introduction. AI and Big Data in Cardiology, Springer International Publishing, pp.1-10, 2023, 10.1007/978-3-031-05071-8_1 . hal-04212037

HAL Id: hal-04212037

<https://hal.science/hal-04212037v1>

Submitted on 25 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Introduction

Dr Andrew King^a

Dr Nicolas Duchateau^{b,c,}*

^a *School of Biomedical Engineering and Imaging Sciences, King's College London, United Kingdom.*

^b *Univ Lyon, Université Claude Bernard Lyon 1, INSA-Lyon, CNRS, Inserm, CREATIS UMR 5220, U1294, F-69621, Lyon, France.*

^c *Institut Universitaire de France (IUF), France.*

** Corresponding author.*

Authors' contribution:

- Main chapter: AK, ND.

Abstract

This chapter acts as an introduction to some of the key terms and concepts that are necessary to understand the subsequent chapters. The terms artificial intelligence (AI), machine learning and deep learning are introduced and their relationship explained. A brief review is provided of some of the key applications of AI in medicine and medical imaging. The role of big data in training effective AI systems is discussed.

Keywords:

artificial intelligence, machine learning, deep learning, big data, data annotation, data curation

Learning Objectives:

At the end of this chapter you should be able to:

- O1.A Explain the relationship between AI, machine learning and deep learning*
- O1.B Give some prominent examples of the application of AI in medicine*
- O1.C Explain the meaning of 'big data' and describe some of its potential benefits and risks*

Aims and Motivation

These days, it seems as if barely a week passes by without a new story in the media about the ways in which artificial intelligence (AI) has changed or will change our lives. From robotics and game playing to self-driving cars and machine translation, AI is projected to hugely impact on our lives in coming years and decades. But without doubt, one of the most promising areas in which AI has the potential to bring benefit to humanity is in healthcare. Projects are underway all over the world that aim to exploit the power of the AI revolution in areas such as drug development, automated diagnosis, treatment optimization and computer-assisted surgery. In a recent Accenture report¹, it was estimated that investment in the use of AI in healthcare would be \$6.6 billion in 2021, with potential benefits to the USA economy alone of \$150 billion by 2026.

Because of this upsurge in interest and investment, a new breed of medical practitioner has begun to emerge in recent years. Whereas once there was little overlap between the clinical and technological specialisms, increasingly researchers and practitioners are occupying a space at the interface between these two exciting fields. The availability of easy-to-use but powerful software tools is enabling doctors to experiment with AI-based technological solutions and to develop their knowledge in this area. Similarly, AI researchers are engaging more and more with clinicians and learning to speak the language of medicine.

This process is blurring the boundaries between the fields of medicine and AI technology and represents a huge opportunity to transform patient care. We believe that this engagement from both sides is an essential step to ensure translation of AI into clinical practice. This book is aimed at people who wish to begin or continue their journey into this cross-over field. We emphasize that this is a journey that will never end. For people beginning their journey from both sides (medicine and AI) there is always more to discover and learn and both fields are developing fast. We hope that this book and the many valuable and insightful contributions of the authors can help readers in this lifelong learning journey.

¹<https://www.accenture.com/sg-en/insight-artificial-intelligence-healthcare>

We have tried to assemble a team of authors who are global leaders in this emerging cross-over field. There is a mix of practising cardiologists, radiologists and AI specialists, but all have a strong interest in seeing their work translated into clinical practice. We have made a deliberate effort to include perspectives from a variety of backgrounds and specialisms to reflect the diversity of the field. None of us are experts in all areas covered in this book, and we believe that we all can and should learn from each other. This book represents a forum for such learning.

What are AI and Machine Learning?

We seem to hear the terms *AI* and *machine learning* a lot these days, whether through reading scientific papers or the popular media. More recently, the term *deep learning* is starting to take centre stage. These terms are often used imprecisely and interchangeably, so we start by introducing their precise meanings.

AI, machine learning and deep learning are strongly related, and are in fact progressively more precise concepts, as illustrated in Figure 1.1. In general, AI refers to the ability of machines (i.e. computers) to exhibit ‘intelligent’ behaviour. Although seemingly a simple definition, in reality what we term ‘intelligent’ is difficult to pin down, and also shifts over time. In what has become known as the ‘AI effect’, many tasks, once they can be effectively performed by computers, are no longer considered to be intelligent. Whatever the definition of intelligence, a wide range of different types of approach can be employed to try to achieve intelligent behaviour. Originally, researchers tried to program intelligence directly, by specifying a series of rules and/or algorithmic steps that, when executed, would lead to the desired behaviour. An example of such an approach is the *expert system*, which is a rule-based inference system that aims to draw conclusions from data given a set of rules that represent expert domain knowledge. Some early attempts were made at using expert systems as clinical decision support tools [11]. For example, the MYCIN expert system [357], developed in the 1970s, was able to recommend antibiotics for patients suffering from severe infections. However, although it worked well, it was never actually used in practice due to technological limitations with deploying it at scale.

The term machine learning was first introduced by Arthur Samuel in the 1950s, who defined it as a “field of study that gives computers the ability to

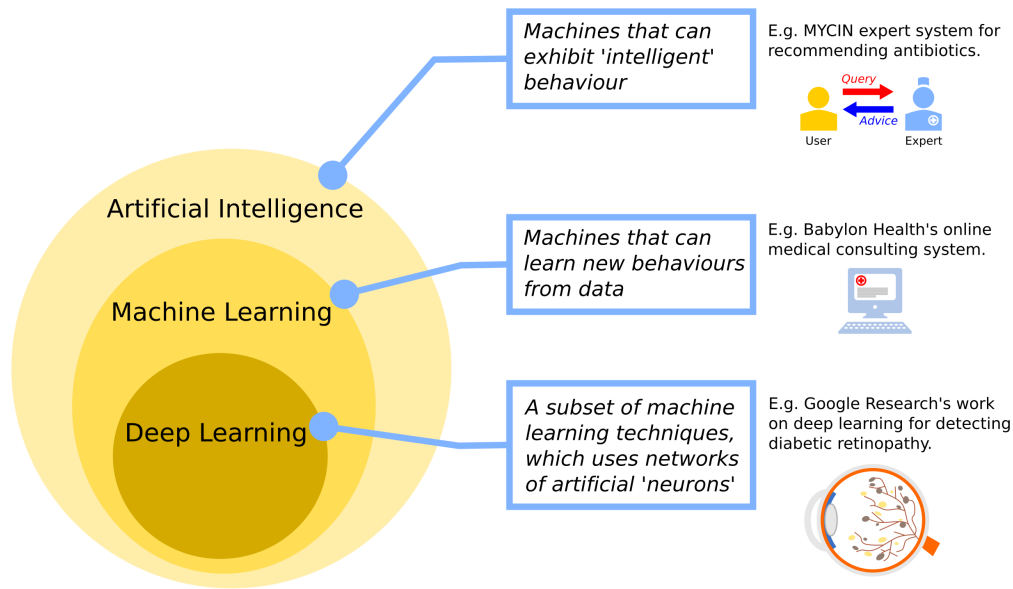


Figure 1.1: The relationship between Artificial Intelligence, Machine Learning and Deep Learning.

learn without being explicitly programmed.” It came to prominence after the traditional AI approach of attempting to directly program intelligence proved hard to achieve. In reality, most tasks that require intelligence are difficult to fully specify in this way. Machine learning tries to overcome this problem by giving the computer the ability to learn new behaviours as it is exposed to new data. Potentially, these behaviours can go beyond what human beings can specify. For example, IBM’s ‘Watson’ machine learning system has been applied to range of applications, including the management of lung cancer². Similarly, Babylon Health³ have developed a machine learning based online consulting system based on knowledge graphs which can learn from user-reported symptoms [327].

Deep learning is closely linked to the idea of *neural networks*, sometimes also called *artificial neural networks*. Neural networks are a specific way of implementing machine learning. They work by defining artificial ‘neurons’, which

²<https://www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-business-in-healthcare/#40cf2fd05402>

³<https://www.babylonhealth.com/ai>

in some sense mimic the operation of neurons in biological systems such as the human brain. These artificial neurons are grouped together into layers, which are connected to form a network. The network as a whole takes a stimulus as input and produces a response as output. For example, consider a simple example in which an AI tool aims to diagnose a disease from a medical image - in this case the image intensities would be the stimulus and the diagnosis would be the response. Deep learning and neural networks are strongly related terms and are sometimes used synonymously. However, deep learning is typically used to refer to neural networks with many layers. Recent hardware and algorithmic developments have made such deep networks possible and they have exhibited great power in many applications, as we will see in Chapter 3. Recent examples of the use of deep learning in medicine include the tool developed by Google Research for the automated detection of diabetic retinopathy and diabetic macular edema in retinal fundus photographs [147] and the tools for automatic contouring of cardiac chambers introduced by, for example, Circle Cardiovascular Imaging⁴.

In this book we use the term AI to refer to any computer-based technique that aims to achieve intelligent behaviour, and we take a fairly broad definition of intelligence. But as we shall see, many of the most promising techniques these days are based on machine learning, and more specifically deep learning.

A Brief History

AI has been around a lot longer than most people realize. Figure 1.2 shows a brief timeline of the major eras and milestones in the history of AI. The term ‘AI’ was first coined at a conference in 1956, and some early successes led to much optimism as to what could be achieved using these new techniques. For example, in 1962 an AI program developed by Arthur Samuel for playing checkers⁵ was able to defeat a master’s checkers player. This early optimism led to some bold, and probably unwise claims. In 1970, Marvin Minsky famously told Life Magazine that “in from three to eight years we will have a machine with the general intelligence of an average human being.” When these claims were not realized, and more investigation revealed some limitations of the techniques being used, interest in AI waned. This led to

⁴<https://www.circlecvi.com/cvi42/cardiac-mri/>

⁵Known as draughts in UK English.

what has become known as the “AI Winter”, in which research funding dried up and little progress was made.

It was not until the 1990s that interest in AI grew significantly again. As computers became more powerful it became more feasible to address many complex problems through ‘brute force’ (i.e. by using lots of processing power), as well as algorithmic sophistication. A famous milestone was reached in 1997 when IBM’s ‘Deep Blue’ AI system beat the world champion Garry Kasparov at chess. This was significant because chess is a massively more complex game than checkers - the ‘state space’ (the number of possible board positions) of chess is approximately 10^{47} compared to 10^{18} for checkers.

This success continued to grow into the new millenium. In 2011, IBM’s question-answering AI program ‘Watson’ won the TV show Jeopardy. Around 2015 there was an explosion of interest in deep learning, prompted by technical developments and also the availability of fast hardware to make the training of very deep networks possible. This led to an acceleration in progress and much interest from large technology companies such as Google, Facebook and Twitter. In 2016 Google’s AlphaGo program beat the 18-time world champion Lee Sedol at Go⁶. This success brought with it public interest and research funding, much of it from industry. This represented a significant shift - for the first time, AI moved beyond a primarily university-based research topic to a tool that many believed could be exploited to solve real-world problems, and hence make money.

Considering this brief history, we can characterize the evolution of AI as starting off with a broad range of approaches, but gradually focusing on more promising specific approaches (i.e. machine learning and now deep learning). This specialization is reflected in this book.

AI in Medicine

The recent success of AI and deep learning has prompted researchers to address many problems in medicine with these techniques, including in cardiology. Figure 1.3 shows the number of papers listed on the medical search engine *Pubmed* that mention the terms ‘artificial intelligence’ and ‘deep learning’. We can see a gradual rise in interest in AI that has accelerated since

⁶Go has a state-space of approximately 10^{170} .

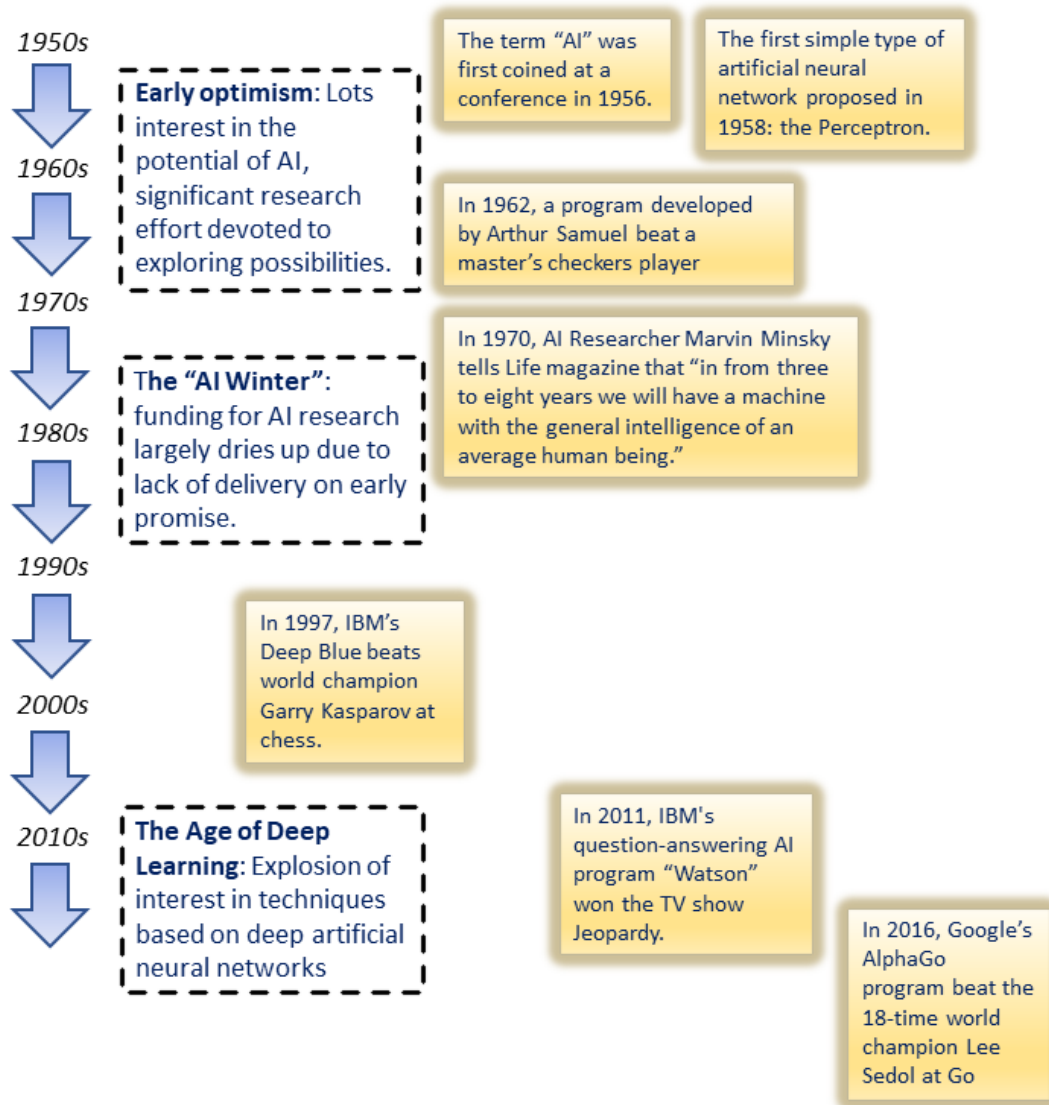


Figure 1.2: A timeline of AI.

the start of the new millennium. But interest in deep learning has grown exponentially in a much shorter time-frame. It can sometimes be hard to “sort the wheat from the chaff” when faced with such a large number of recent publications, so below we have tried to pick out some of the more significant recent advances in the use of AI in medicine⁷.

A prerequisite to the development of many of the most powerful AI models is the formation of large scale databases. A significant milestone here was the commencement of the UK Biobank project⁸ in 2006. This is a 30-year programme to collect genetics, imaging and other data from half a million participants, and to make the data available to researchers all over the world. This project has so far led to more than 1,000 research publications, as well as inspiring other similar initiatives. Such databases have enabled machine learning models to be trained using much larger amounts of data than was previously feasible. In part because of this, it has recently become possible for AI to approach and even match the level of human expert performance in some applications. For example, in 2016 expert performance was achieved in the detection of diabetic retinopathy [147] and in 2017 in the identification of cancerous skin lesions [117]. These, and other achievements have prompted an increase in interest in commercialization of AI in medicine. A milestone here was achieved in 2018 when HeartFlow’s ‘FFR_{CT}’ analysis software was approved for reimbursement in the USA healthcare system⁹. This tool allows clinicians to noninvasively determine the impact of artery narrowing on blood flow to the heart, and hence better select patients for intervention. It is likely that coming years will see further commercial exploitation of AI. Other promising recent research includes the prediction of the onset of disease from electronic health records [270], survival prediction based on motion extracted from cardiac magnetic resonance (CMR) images [47], automation of cardiac functional analysis from CMR [335] and breast cancer screening from mammography data [260, 419].

⁷Any such summary will necessarily be selective and subjective to a degree, and we apologize and take responsibility for any omissions. There are many good papers that we have not been able to include here.

⁸<https://www.ukbiobank.ac.uk/>

⁹<https://www.heartflow.com/reimbursement-resources>

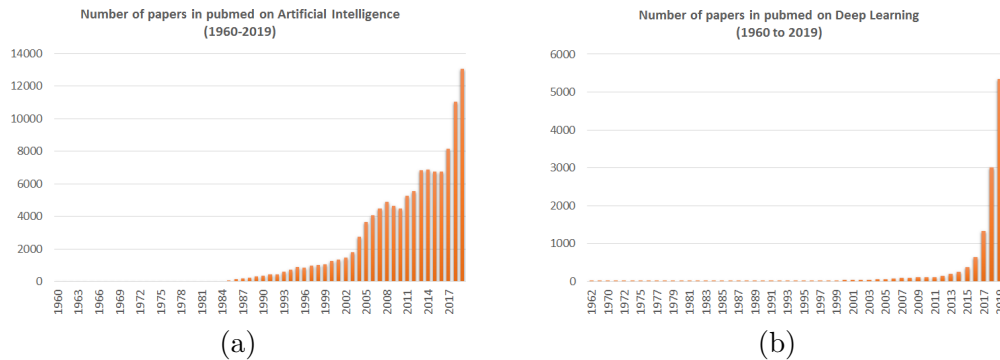


Figure 1.3: The number of papers between 1960 and 2019 on the *Pubmed* search engine on the subjects of (a) Artificial Intelligence, and (b) Deep Learning.

The Role of Big Data

As noted in the previous section, many of the most powerful machine learning and deep learning techniques rely on large amounts of data to train and evaluate their performance. One of the reasons for the recent upsurge in interest in such techniques is the growing availability of large scale public databases. The term *big data* is often used to refer to such databases. There is no hard-and-fast definition for what constitutes big data, but typically these days several thousands of datasets are available for many medical applications. As well as a large number of datasets (i.e. individuals), big data refers to the gathering of many different variables from each individual. For example, the UK Biobank project, mentioned above, plans to acquire a range of lifestyle, nutritional and medical data from half a million volunteers, as well as MR imaging data (brain, heart, abdomen, bones and carotid artery) from 100,000 volunteers. For cardiology and in particular cardiac imaging, such initiatives pave the way for a new era of very large databases, compared to the previous generation of imaging studies which typically comprised hundreds to thousands of subjects [375].

As an aside, it is important to be precise when discussing the sizes of databases. For example, the data from a $100 \times 100 \times 100$ voxel MR scan acquired from a patient could be described as consisting of one million voxel intensity values, or one hundred slices, or just data from a single patient. How we count sizes of databases depends upon what we want to do with them. For example, if

our task is to diagnose a disease, then the units of interest are patients, and so the size of the database should be based upon the number of patients. Alternatively, if our task is to segment¹⁰ a two-dimensional image, then the units of interest will be image slices.

Much of the value of big data comes from *annotations*. Annotation refers to adding information to, or ‘labelling’ the data. For example, to return to our example of diagnosing a disease from a medical image, annotation in this case would involve adding information about disease diagnosis to the images used for training the AI tool. Alternatively, for a segmentation task annotations would consist of ‘ground truth’ segmentations of certain structures. What we choose to annotate depends upon the task we want to address. But most tasks involve some form of data annotation to create a useful database for training our AI model. The process of annotation is normally at least partly manual, and so can be quite time-consuming. Gathering, organising and annotating data has become an important activity in AI. Collectively, this process is known as *data curation*.

Of course, the benefits of big data come with associated risks, as the new data may not always come with new knowledge [127]. Smaller and well-controlled databases often consist of a limited number of variables that are more specific to the application of interest, meaning that they already focus on the question that we wish to address. They also come with better ‘fidelity’ of the data. By data ‘fidelity’ we refer to the presence of all data fields for all individuals and the absence of corruption or errors in their values. For example, high fidelity data for our disease diagnosis example would mean having high quality images with no serious artefacts and diagnosis annotations without errors for all images. With big data, ensuring data fidelity becomes much harder and requires automated techniques for controlling the quality of data and annotations, as we will see in Chapter 7. Collecting more samples and variables also means adopting a broader view on the studied diseases/applications, with potential redundancy between variables, lower or unknown relevance of some variables to the disease/application, and higher demands in storage and computational power. A global roadmap for balancing these risks with the potential for new discoveries would be of high value, as promoted by the neuroscience community [127]. In the field of cardiology, some

¹⁰Segmentation refers to the process of outlining, or delineating, structures of interest in images, such as organs or tumours.

of these practical issues will be specifically discussed in Chapter 9.

Outlook

This chapter has provided an overview of AI and machine learning in medicine, deliberately avoiding any technical details. But it is clear that we are living in interesting and exciting times, and the opportunities for turning technological advances into patient benefit are huge. The use of AI in medicine is a fast-moving field - a book like this written in 10 or 20 years' time would almost certainly look very different. But realising the potential benefits of the AI revolution in medicine will require hard work and commitment - a commitment to learning and expanding our knowledge into unfamiliar areas. The next chapter is intended to help readers in this learning journey by reviewing some more technical aspects of AI and machine learning in an accessible way.

Acknowledgements

AK was supported by the EPSRC (EP/P001009/1), the Wellcome/EPSRC Centre for Medical Engineering at the School of Biomedical Engineering and Imaging Sciences, King's College London (WT 203148/Z/16/Z) and the UKRI London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare.

ND was supported by the French ANR (LABEX PRIMES of Univ. Lyon [ANR-11-LABX-0063] within the program "Investissements d'Avenir" [ANR-11-IDEX-0007], and the JCJC project "MIC-MAC" [ANR-19-CE45-0005]).