EMERGENCY BENEFITS AND RISKS OF ARTIFICIAL INTELLIGENCE

Thomas V. Robertson
Thinking Teams[1]

thinkingteams@gmail.com

Abstract

The research and development field of Artificial Intelligence develops computer systems (called “AIs”) that can perform tasks normally associated with human intelligence. Fifty years ago, AIs could play excellent checkers, solve algebra word problems, and prove logic theorems. In recent years AI capabilities have increased rapidly, taking advantage of exponentially increasing computational power and machine learning to identify patterns by sifting through massive amounts of data available online and in digital form. Newly developed AIs do a good job of translating spoken or written languages, recognizing faces, driving cars, and beating the best human players in the games of chess and go.

AI has been beneficially applied to emergency management, using historical and recently collected data to predict the impacts of earthquakes, flooding, and crop diseases. Our online, “wired” world creates mountains of digital data that emergency managers can mine using AI to assess emergency situations and plan and prepare for responses.

However, AI can create new risks to our societies, and according to some experts, potentially new forms of disasters and emergencies. Today, use of AI threatens cybersecurity and privacy, and there is a growing debate about the safety of autonomous systems. Scientific and technical experts such as Stephen Hawking and Elon Musk have warned that as AIs become more intelligent and more ubiquitous, they might become dangerous masters of inferior humans. This paper reviews these concerns, their plausibility, and the role emergency management might take to reduce the risks associated with AI.

Keywords: Artificial Intelligence, Emergency Management, Deep Learning, Super Intelligence, AI Risks

Introduction

For over 60 years researchers in the field of Artificial Intelligence (AI) have been working to create “smart” computer systems that can solve problems and perform tasks usually associated with human intelligence. Although along the way there have been notable accomplishments, as well as setbacks, we find ourselves today with artificially intelligent machines that translate languages, drive cars, beat human game champions, and perform other impressive feats. Perhaps even more amazing, these machines have shown a remarkable ability to learn and improve their performance, even to the point of outperforming their creators (Silver et al., 2017).

Artificial Intelligence has increasingly been applied to emergency management (GRDRR, 2018), drawing on its ability to identify patterns in imagery and large digital data sets to assess vulnerabilities and predict outcomes. Internet connectivity and the availability of smartphone apps makes many of these capabilities accessible to local and regional emergency managers. This paper will review examples of such applications of AI.

The emergence of AI brings with it more than useful tools. Because AI is becoming more pervasive and more capable, it is changing the way we live in ways that impact emergency management. Near term, there are new threats from autonomous terrorist weapons and social manipulation, as well as new vulnerabilities in our cyber connected/enabled infrastructure (Knight and Hao, 2019). Longer term, there are concerns that AI could become so capable that it could become a treat to humankind (Tegmark, 2017). This paper explores these threats, and the role Emergency Managers might play to minimize the risks to our societies.

[1] 13220 NW 33rd Ave, Vancouver, WA 98685 USA

What is Artificial Intelligence?

The research and development field of Artificial Intelligence develops computer systems (called “AIs”) that can perform tasks normally associated with human intelligence. Researchers have approached this challenge from several angles (Russell and Norvig, 2015), that have emphasized both “***thinking***” – learning and problem solving as humans do, for example solving puzzles, and “***doing***” – carry out activities in the physical world, such as a robot vacuum cleaner. These two aspects are closely related and often combined, for example when a rescue robot plans its path through a building. The “thinking” aspect can be thought of as primarily producing information, and the “doing” aspect produces physical effects in the world. Of course, information and physical effects are both important to emergency management. Figure 1 depicts modern examples of AI: clockwise from the far left - machine translation, digital assistant, self-driving car, Go game champion, and aerial drone.

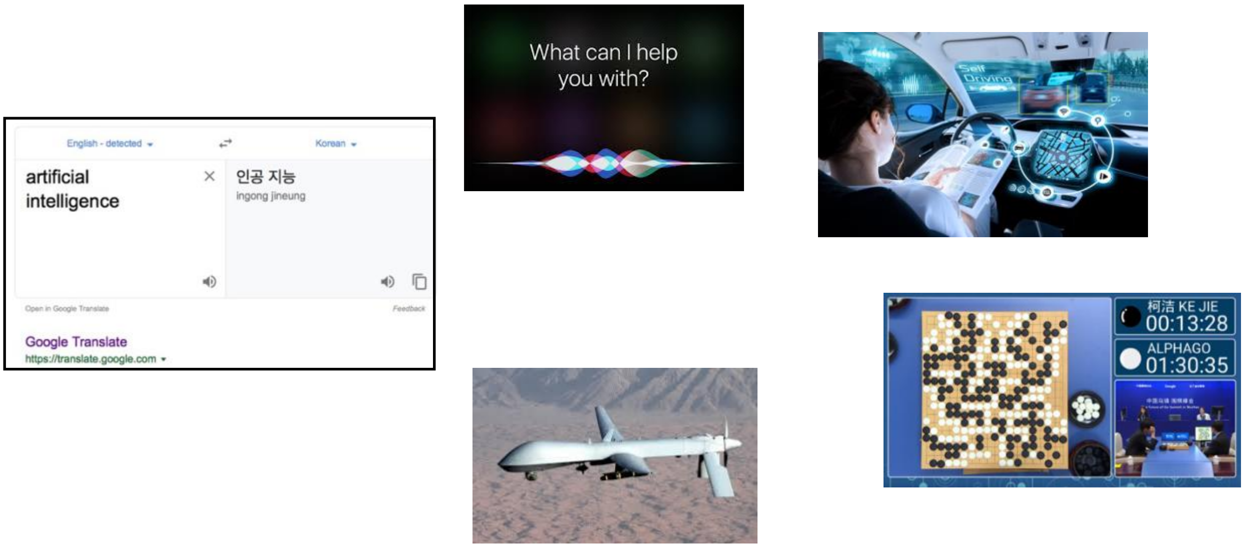
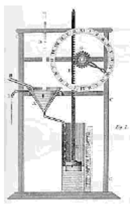


Figure 1: Modern applications of Artificial Intelligence

AI has a long history and a review of some of its milestones can help illuminate the nature of its capabilities and limitations.

The roots of AI can be traced back to the first self-controlling machine, a water clock developed by Ktesibios of Alexandria around 250 B.C. This machine solved the problem of variable flow rate as the water reservoir emptied and was the first time a non-living entity was able to adapt its behavior to changes in the environment (Russell and Norvig, 2015).



In 1951 Claude Shannon, founder of Information Theory (Collins, 2002), developed one of the first learning machines – a life-size mouse controlled by magnetic relay circuits adapted from telephone circuits. The mouse was able to learn to traverse a 25-square maze from any point in the maze. The maze could be modified, and the mouse would relearn the portions of the maze that had changed (Klein, 2018).

The far more complicated domain of the checkers board game was tackled by Arthur Samuel in 1955 (Samuel, 1959). The program pioneered the use of reinforcement learning and improved its game by playing itself, techniques important to today’s AIs. Samuel’s program was not a champion, however it was an early counter example to “computers can only do what you program them to do”, since it was able to beat its creator at checkers.

A different sort of AI was popular in the 1980s – Expert Systems. These systems could make logical inferences and use knowledge stored in databases. An example was the R1 system used by Digital Equipment Corporation (DEC) to configure computer systems for customers (McDermott, 1982). DEC estimated that R1 saved them $25M per year in operating costs, much of it because the configurations shipped to customers were correct the first time, reducing the need for returns and reshipment.

Starting in the late 1990s, game-playing AIs began making the news, as advances in computing speed and memory allowed improved game performance based upon techniques developed in prior years. In 1997 IBM's Deep Blue defeated champion Garry Kasparov in a 3.5 to 2.5 match (McPhee et al., 2015). In 2011 IBM's Watson defeated two of Jeopardy's greatest Champions (Best, 2013). Finally, in 2015 AlphaGo, a program developed by DeepMind Technologies became the first computer program to beat a human professional player. This program was followed by an even more powerful version AlphaGo Zero in 2017, which was able to beat the original AlphaGo using much less computer power. AlphaGo Zero also distinguished itself by not relying on knowledge of previous games played by humans, instead learning winning moves only through trial and error (Silver, 2017).

In addition to progress in "thinking" AI game players, researchers have made significant advances in "doer" robots. An example was the winner of the Defense Advanced Research Projects Agency (DARPA) 2005 Grand Challenge. The Grand Challenge was created to spur innovation in unmanned ground vehicle navigation and offered a prize of $2M to successfully navigate a 142-mile course through the Mojave Desert in southwestern US. The prize was won by Stanford University's "Stanley" vehicle, a Volkswagen Touareg outfitted with sensors, actuators, and computing platforms that navigated the course in 6 hours 53 minutes (Thrun et al., 2007).

The AI achievements cited above are founded on the development of a variety of AI technologies, and on the rapid advance of computing technologies in general. Again, it is useful to review these underlying technologies, to give a sense of the weaknesses and strengths of AI.

For the purpose of this review, AI technologies are categorized as four types:

- Neural networks
- Symbolic logic systems
- Knowledge-based systems
- Statistical methods.

Neural networks were one of the earliest forms of AI, and after falling out of favor for some years, they are now at the center of modern AI breakthroughs. Neural networks are inspired by human brain physiology (Figure 2).

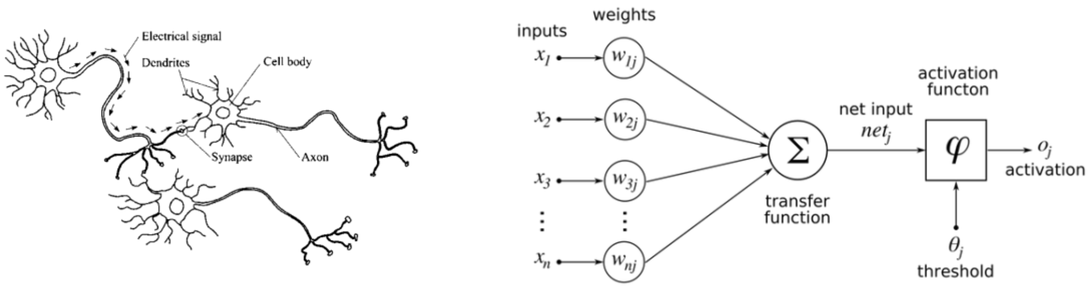


Figure 2: Natural and Artificial Neural Networks

In living brains, neurons are interconnected through synapses that allow neurons to “fire” in response to stimulation from other connected neurons. It is known that these interactions among neurons is the source of mental and nervous activity in the brain, and early researchers sought to imitate natural neural networks with simplified models called artificial neural networks, such as the one shown in Figure 2. In these artificial networks, a set of inputs (e.g., representing measurements from a photo) is presented to an “input layer” in which each input is multiplied by a “weight”. The inputs times weights are then summed and if the sum exceeds a threshold, an “activation” signal is output. The basic idea is that this mimics a natural neuron that will fire when certain other neurons excite it. A foundational work by Warren McCulloch and Walter Pitts (McCulloch and Pitts, 1943) showed that any computable function could be implemented by some network of connected neurons, and that all logical functions (and, or, not, etc.) could be implemented by simple neural networks. Donald Hebb (Hebb, 1949) demonstrated a simple way to modify artificial neuron weights to allow the network to “learn”.

Early AI neural networks caused excitement and much speculation about thinking machines, however early work failed to produce impressive capability. However, starting in the late 1980s advancing computer power and a better understanding of training complicated networks led to steady progress, so that now some of the most impressive AI feats are performed by very large neural networks. For example, AlphaGo Zero has 361 inputs and about 40 layers of weights.

The second AI technology, the symbolic logic approach, also has early roots. Whereas neural network approaches to AI attempt to mimic physical brain function, symbolic logic approaches focus on the sort of reasoning humans carry out in conscious thought. Symbolic logic follows in the tradition of the Greek philosopher Aristotle by formally expressing rules such as “All mammals are animals. John is a mammal. Therefore, John is an animal.” AIs that use symbolic logic can make inferences and solve problems based on such logical rules. In 1956 Allen Newell and Herbert Simon created a program called the Logical Theorist (Newell et al., 1957), which was able to prove theorems in symbolic logic.

Knowledge-based systems add knowledge to the rules of logic, knowledge about what are the elements that make up our world (or a particular domain of interest, like DEC computer configurations), and how they are related. It is clear that humans learn fundamental facts about their world very young, and that this knowledge is critical to making sense of language and every situation encountered in the world. Early work in limited “worlds” such as the “blocks world” (Minsky, 1968) demonstrated how to encode knowledge and use it to solve problems such as that encountered by a robot trying to reconfigure blocks on a desktop.

Early work in knowledge-based systems laid out a basic approach that seemed very general, and it was hoped that powerful AIs could be built by using this approach with a sufficient quantity of programmed-in facts. The Japanese Fifth Generation Computer Systems Project, launched as a 10-year initiative in 1982, was representative of the hopes for symbolic logic-based AIs. Japan believed that by developing fast parallel logic computers and large databases, they could become the world leaders in computing technology. Their goal was credible enough that the rest of the world took notice, and rival initiatives were started in the US, UK, and Europe. Unfortunately, the expected major breakthroughs never occurred (Pollack, 1992).

The fourth AI technology, statistical methods, represents acceptance into AI of techniques that predated AI. Early investigators of AI ignored these methods and sought breakthroughs by pursuing techniques they viewed as fundamentally different from the statistical methods, which had been developed in areas such as control theory and statistical pattern recognition. An early example of this work was the groundbreaking work of Norbert Weiner (Weiner, 1948), which suggested the possibility of intelligent machines based on feedback control systems. More recently, AI has recognized the power of rigorous statistical methods and embraced techniques such as hidden Markov models for use in areas such as speech recognition.

Characteristics of the four types of AI technology are summarized in Table 1.

Table 1: Summary of Four AI Technologies

AI Technology	Inspiration	Roots	Modern Use
Neural Networks	Brain physiology	Neuron model	Deep learning neural networks
Symbolic logic	Human reasoning	Aristotle	Formal languages
Knowledge Based Systems	Common sense/expertise	Blocks world	Expert systems
Statistical Methods	Mathematical rigor	Cybernetics	Speech recognition

Although each of these technologies has been pursued by various researchers as a singular approach to AI, successful AI systems today are engineered to incorporate combinations of these technologies. Even the self-learning deep neural networks celebrated today rely on clever engineering of multi-element systems, belying the claim that these AIs have designed themselves.

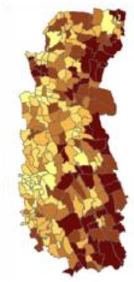
How can AI Help Emergency Managers?

Many of the tasks faced by emergency managers can be assisted by AI-based tools, for example, AI can be used to (GFDDR, 2018):

- Analyze community and regional data, as well as remotely sensed data, to assess vulnerabilities and risks associated with structures, population patterns, poverty, transportation and communication networks, etc.
- Analyze historical patterns to predict disaster impacts from flooding, earthquakes, landslides, etc.
- Monitor real-time data and generate alerts for potential earthquakes, wildfires, etc.
- Perform post-disaster damage assessment.

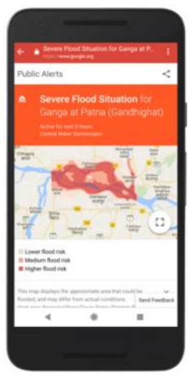
A few specific examples are now presented.

<u>Urban Poverty Mapping</u>



Poor urban areas are especially vulnerable to disasters and poverty data is in scare supply and difficult to collect. Researchers at Oak Ridge National Laboratory in the US have developed a AI-based technique to identify poor, informal settlements from high-resolution satellite imagery (Graesser et al., 2012). Their approach used a variety of spatial, structural, and contextual features to classify areas as formal, informal, and non-settlement classes. The method was tested in Caracas, Kabul, Kandahar, and La Paz, and demonstrated that good accuracy could be obtained using the same features in these diverse areas.

<u>Predicting Floods in India</u>



Twenty percent of global flood fatalities occur in India. Google has been partnering with India's Central Water Commission to develop AI-enabled flood forecasting and early warning (Matias, 2018). Google uses a variety of elements such as historical events, river level readings, terrain and elevation, to run hundreds of thousands of simulations for each location to create river flood forecasting models that can more accurately predict where and when a flood might occur, and also how severe it will be.

Wildfire Prediction in California

In California, two high school students invented a device to predict the probability of a forest fire occurring (Shah, 2018). The device is placed in the forest and takes real-time photographs and measurements of humidity, temperature, carbon monoxide/dioxide, and wind. This data is then used with a deep learning neural network to predict the probability of a fire.

Fighting Fall Armyworm in Africa

Fall Armyworm has infected millions of hectares of maize in Africa, threatening the food security of more than 300 million people. The Food and Agricultural Organization of the United Nations has launched the Fall Armyworm Monitoring and Early Warning System (FAMEWS), a mobile app that is used to collect data from the field and advise farmers (UNFAO, 2019). The app was initially implemented in Madagascar and Zambia, and later rolled out across all the countries in Sub-Saharan Africa affected by the insect.

How will Artificial Intelligence Impact Emergencies?

Clearly, AI can provide tools that will be useful to many facets of society, including Emergency Managers. However, beyond providing tools that fit various niches in today's society, AI has the ability to shape society itself. Because Emergency Mangers seek to protect the welfare of society, risks to societies brought about by AI are of direct concern to Emergency Managers.

Today, we are seeing the beginnings of threats to society brought about by the increasing capability and prevalence of AI. These threats will impact Emergency Management. Examples include:

- **Autonomous weapons** – weapons such as drones and autonomous vehicles present new dangers in the hands of criminals and terrorists, expanding the range of possible emergency incidents.
- **Smart infrastructure** – the internet of things, big data, and pervasive wireless communication are leading to more highly automated buildings, cities and regional infrastructure. This brings many benefits, of course, however it also opens new vulnerabilities to catastrophic failures and cybercrime.
- **Social manipulation** – AI "bots" on social media are being used to influence the public, in some cases creating conflict and violence, and in other cases resistance to emergency management efforts.
- **Automation and loss of jobs** – many blue-collar and white-collar administrative jobs are being lost to automation. Although AI and automation create the need for more engineers and specialists, these jobs are not as broadly accessible to typical populations as the jobs being eliminated by automation. There is growing concern that this will create more income inequality in some parts of the world, and change patterns of poverty. This is of concern to Emergency Managers, since if affects the distribution of vulnerability in their communities.

Longer term, where will AI take us? There is much speculation about this issue. One school of thought envisions the progression of AI first to Artificial General Intelligence, equivalent to human intelligence, then to Superintelligence, by virtue of AI's ability to self-improve. Once AIs are superintelligent, the speculation continues, humankind will essentially be at the mercy of these superior beings, who may well decide against humans in favor of their own goals. This may sound fanciful, however prominent figures such as Stephen Hawking, Elon Musk, and Bill Gates have expressed strong concern that AI might eventually lead to disaster (Parkin, 2015).

A somewhat contrary school of thought challenges the likelihood that superintelligent AIs will be an existential threat to humankind at any time in the foreseeable future. This view paints the fear of superintelligent AIs as misguided, because

- Even the most capable AIs developed so far are quite narrow in their intelligence, compared to even a human child. As impressive as an AI Go master is, the "universe" "understood" by that AI is a 19 x19 game board, black and white stones, and the simple Go rules. There is no comparison of the complexity of that "world" compared to world as understood by even a human child. Computers can do arithmetic billions of times faster than humans. Does that make us fear their dominance over humans? (Though computers can certainly replace clerks!)
- It is true that there is "machine learning", and its existence leads some to believe AIs could use this inexhaustibly to become superintelligent. However, the learning performed by machines is not comparable to the learning capabilities of humans. First of all, machines learn within a framework heavily structured by human engineers, e.g. arrays of weights that are adjusted as neural networks are trained. AlphaGo Zero learned to play excellent Go, but it did not learn the framework necessary for it to learn. That framework had to be provided by human engineers. A second contrast between machine and human learning is the vast data sets needed for machine learning. Human children learn language by hearing relatively few works compared to the millions or billions used to train learning machines. This is not only a matter of efficiency, but a fundamental limitation as well. Machine learning is limited to those domains where AIs can learn patterns by being shown very large numbers of examples. There are many parts of the real world that do not fit that scenario.
- Agency – getting things done in the world – is greatly facilitated by intelligence. In fact, it could be said that humans with the highest intelligence have the greatest potential impact on the world, for good or evil. However, there is so much more to agency than intelligence, things like motives, physical power, collaboration, physical possibilities and constraints. Even if an AI could achieve a high level of "intelligence" along some dimension, this would be far from guaranteeing it the power of agency in the real world.

Regardless of the likelihood of existential threats to humankind from AI, there is no question that there are risks as well as benefits associated with AI. How can Emergency Managers prepare to deal with these risks?

How Can Emergency Managers Prepare for Artificial Intelligence?

Emergency Managers need to be alert to the opportunities and risks of AI. Here are some of the questions Emergency Managers can keep in mind as the use of AI continues to permeate our societies:

- **Mitigation** – What opportunities and threats are presented by more highly interconnected and automated infrastructures? How might back-up plans mitigate risk?
- **Prevention** – What new safety standards are needed for AI? How might these be audited?
- **Response** – Do emergency management AIs need to be built to respond to emergencies in the age of AI? To counteract "runaway" or failing AIs?
- **Recovery** – Do Emergency Managers become the keepers of older technology infrastructure for recovery and backup?

Emergency Managers have the opportunity to be among those taking the initiative to explore AI risk and safety issues. These initiatives might be taken within existing Emergency Management Agencies and Organizations, or Emergency Managers might collaborate with existing organizations concerned with AI risks and safety, such as the Future of Life Institute (FLI, 2019) and the Future of Humanity Institute (FHI, 2019).

Conclusion

For the past 60 years, the field of Artificial Intelligence has worked to develop computer systems that solve problems and perform tasks usually associated with human intelligence. Progress along the way has captured the imaginations of researchers and the public, with imagination often out-pacing actual AI capability. In spite of frequent setbacks, when hoped-for achievements proved much more elusive than expected, much progress has been made. Today we have AIs that are game-playing champions, self-driving vehicles, language translators, and personal assistants.

A number of AI tools have been developed to aid Emergency Managers, and more are being developed. Some of these are as convenient as a smartphone app, and help assess risks, predict wildfires, flooding, and other disasters, and support pooling knowledge to fight crop disease.

AI also brings with it risks. The nature of the emergencies we face in the modern era are impacted by autonomous weapons, smart infrastructure, social manipulation, and job automation, all enabled by AI. In the longer term, as AIs become even more capable, we may find ourselves at odds with the AIs we have created, bringing risks and security issues that will challenge Emergency Managers. Emergency Managers are encouraged to explore the potential benefits of AI in their work, and to recognize their role as a profession in the safety and security of a world shared with AIs.

References

Best, J. (2013). IBM Watson: The Inside Story of How the Jeopardy-Winning Supercomputer Was Born, And What It Wants To Do Next. Tech Republic, CBS Interactive, New York, NY USA https://www.techrepublic.com

Collins, G. (2002). Claude E. Shannon: Founder of Information Theory. Scientific American, New York, NY USA , https://www.scientificamerican.com/article/claude-e-shannon-founder/

FHI (2019). Future of Humanity Institute, University of Oxford, https://www.fhi.ox.ac.uk

FLI (2019). Future of Life Institute, https://futureoflife.org

GFDDR (2018). Machine Learning for Disaster Risk Management. World Bank Global Facility for Disaster Risk Reduction (GFDRR) Guidance Note, Washington, D.C. USA

Graesser, J. et al. (2012). Image Based Characterization of Formal and Informal Neighborhoods in an Urban Landscape. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 5, No. 4

Hebb, D. (1949). The Organization of Behavior, Wiley, Hoboken, NJ, USA.

Klein, D. (2018). Mighty Mouse. MIT Technology Review, Cambridge, MA, USA, https://www.technologyreview.com

Knight, W. and Hao, K. (2019). Never Mind Killer Robots – Here are Six Real AI Dangers to Watch Out for in 2019. MIT Technology Review, Cambridge, MA, USA, https://www.technologyreview.com

Matias, Y. (2018). Keeping People Safe with AI-Enabled Flood Forecasting

McCulloch, W. and Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. The Bulletin of Mathematical Biophysics, Vol. 5, No. 4, pp. 115 – 133

McDermott, J. (1982). R1: A Rule-Based Configurer of Computer Systems. Artificial Intelligence, Vol. 19, No. 1, pp. 39 – 88, Elsevier, Atlanta, GA USA

McPhee, K. et al. (2015). Deep Blue, IBM's Supercomputer, Defeats Chess Champion Garry Kasparov in 1997. New York Daily News, New York, NY USA https://www.nydailynews.com

Minsky, M. ed (1968). Semantic Information Processing, The Massachusetts Institute of Technology, Cambridge, MA USA

Newell, A. et al. (1957). Empirical Explorations with the Logic Theory Machine. Proc. Western Joint Computer Conference, Vol 15, pp 218 -239

Parkin, S. (2015). Science Fiction No more? Channel 4's Humans and Our Rogue AI Obsessions, The Guardian, London, UK https://www.theguardian.com/tv-and-radio/2015/jun/14/science-fiction-no-more-humans-tv-artificial-intelligence

Pollack, A. (1992). 'Fifth Generation' Became Japan's Lost Generation. The New York Times, June 5, 1992, Section D, Page 1. https://www.nytimes.com/1992/06/05/business/fifth-generation-became-japan-s-lost-generation.html

Russell, S. and Norvig, P. (2015). Artificial Intelligence – A Modern Approach 3rd Edition, Prentice Hall, London, UK

Samuel, A. (1959). Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, Vol. 3, No. 3, pp. 211 – 229.

Shah, A. (2018). Fighting Fire with Machine Learning: Two Students use TensorFlow to Predict Wildfires, Google Blog, https://www.blog.google/technology/ai/fighting-fire-machine-learning-two-students-use-tensorflow-predict-wildfires/

Sliver, D., et al. (2017). Mastering the Game of Go Without Human Knowledge. Nature, Vol. 550, pp. 354 – 371, Macmillan, London, UK

Tegmark, M. (2017). Life 3.0 Being Human in the Age of Artificial Intelligence, First Vintage Books, New York, USA

Thrun, S. et al. (2007). Stanley: The Robot That Won the DARPA Grand Challenge. In: Buehler M., et al. (eds) The 2005 DARPA Grand Challenge. Springer Tracts in Advanced Robotics, vol 36. Springer, Berlin, Heidelberg

UNFAO (2019). FAO Launches Mobile Application to Support Fight Against Fall Armyworm in Africa, https://www.un.org/africarenewal/news/fao-launches-mobile-application-support-fight-against-fall-armyworm-africa

Weiner, N. (1948). Cybernetics, Wiley, Hoboken, NJ, USA.