Automation and Collaborative Robotics

A Guide to the Future of Work

Peter Matthews
Steven Greenspan, PhD

Apress®

Automation and Collaborative Robotics

A Guide to the Future of Work

Peter Matthews
Steven Greenspan, PhD

Apress®
2020

Table of Contents

About the Authors

Peter Matthews, based in the UK, is a writer and research scientist. Peter has more than 40 years of IT experience ranging from mainframe/Unix programming, development and relational databases to secure cloud computing, DevOps and cobotics. Peter's research work has been concerned with leading edge technology for a major proportion of that time. Projects have included machine learning algorithms for soccer clubs, multi-level secure database systems and object-oriented application infrastructures. He has also led groups investigating the influence of macro social, political and economic trends on future technology.

Peter's current research is focused on automation and robotics. Past research has covered Internet of Things, cloud computing, data curation, smart buildings and cobotics. Peter has been the CA project lead on cobotics under the auspices of the Centre for Visual and Decision Informatics, a USA National Science Foundation initiative.

Peter has authored and co-authored academic papers including "Data Is the New Currency" in the proceedings of New Security Paradigms Workshop. Other writing includes books on "Ingres Visual Programming", co-authoring of "The Innovative CIO: How IT Leaders Can Drive Business Innovation" and co-editor/contributor to "MOdel-Driven Approach for design and execution of applications on multiple Clouds".

Steven Greenspan, PhD is an innovator of information and communications technology (ICT), with over 50 publications in peer-reviewed journals and 79 US patents, including the first patent to describe two-factor and two-device authentication and authorization. This invention is widely used throughout the world to ensure secure access to web services and it showcased the value of keeping the human-in-the-loop. His current research interests include user experience and collaboration in complex systems, differential privacy, ethical decision-making, and innovative approaches for integrating scientific research into socially responsible applications and services. Outside of his professional work, he devotes much of his time to community groups that focus on environmental, economic and social justice. He also serves on the advisory boards of the AABGU/ Philadelphia Academic Bridge, and several startups.

Steve has a PhD in Cognitive Psychology from the State University of New York at Buffalo and conducted postdoctoral research at the University of California at San Diego and Indiana University. During the writing of this book, he was a Visiting Scholar at the University of Virginia. Previously, he was a Research Scientist and Vice President of Strategic Research at CA Technologies managing an international team of information technology scientists. He has also consulted to Avaya on the UX design of mobile phones, and was a Distinguished Member of the Technical Staff at AT&T Bell Laboratories.

About the Technical Reviewer

George Watt became passionate about technology and innovation at a very early age and built his first "computer" out of cardboard boxes somewhere around age 5. George led the design workshops for the accelerator program described in this book and created and deployed its foundation artifacts and ceremonies. Throughout his career, George has delivered innovations of his own, such as a knowledge base for a neural network–based predictive performance management solution, one of the earliest private clouds (2005), and a lightweight event management agent. A transformative leader, George has spearheaded initiatives that have enabled organizations to address complex technology problems, deliver new business benefits, and drive millions of dollars in savings and productivity gains. George began his technical career as a systems programmer/sysadmin and systems engineer. He has held many national and global leadership positions and has led global teams spanning North America, Europe, Asia, and Australia. George is co-author of *The Innovative CIO* (Apress, 2012) and *Lean Entrepreneurship* (Apress, 2019) and tweets @GeorgeDWatt.

Acknowledgments

First, as always, I would like to thank my wife Pat and my daughter Georgina for their patience and help. They have put up with my incessant questions as well as keeping on top of the catering. Their help has been invaluable; without it this book would not have seen the light of day.

I owe a debt of gratitude to Karen Sleeth for her encouragement and support preparing the book proposal and to my three interlocutors who were interviewed or responded to emailed questions. The late Serge Mankovski has been a valued colleague for more than 10 years, and his innovative mind and big picture thinking has been instrumental in both automation and collaborative robotics sections of this book. Arnor Solberg has been unstinting in helping me understand healthcare and IoT as well as strategies for data management and data fusion. Moncef Gabbouj is a leading thinker in signals processing and the research he and his students are conducting into new AI models for processing complex data and fused data. All three have helped enormously in developing some of the ideas expressed here.

I would also like to thank Steve Greenspan, my co-author, who has as always given new perspectives in any discussion. It has been great working with him and finally getting this book ready.

—PM

Writing a book is a long journey filled with frustration and joy, confusion and insight. I am thankful to my wife and family for their encouragement and support. Every day, they bring joy. I am especially grateful to my wife, Wendy, and son, Jonathan, who read through one or more chapters and provided me with provocative questions, suggestions, and wise judgment. Rodney Wallace, an economist and author of the book

Drowning in Potential, read an early version of Chapter 1. I am grateful for his advice and for helping me avoid certain errors. The remaining errors are of course my own.

My understanding of information technology, machine learning, and social science was influenced by many outstanding teachers, colleagues, and friends. In particular for this book, I am grateful for the many conversations about machine learning, wearable computing, and robotics that I had with Steve Crandall, Deborah Danielson, George Watt, and the late Serge Mankovski. Serge and I co-authored many papers and patents, and I am thankful for the creative energy that enlivened our conversations about IoT, machine learning, and automation. And of course, I am grateful to my co-author, Peter, for his friendship, for his perspective, and for convincing me to join him on this journey. He was a delightful and inspiring companion.

—SG

We acknowledge the encouragement and feedback we received from George Watt, our former manager at CA Technologies. He was a very thorough technical reviewer, and the book is much better for his questions, insights, and suggestions.

We also acknowledge the contribution of Maria Velez-Rojas who led many of the discussions on data and visualization and shared her expertise but for personal reasons was not able to contribute as a co-author.

Lastly, we thank Rita Fernando Kim for her patience, sage advice, and keeping us on track. We also thank Susan McDermott and the Apress team for their patience and guidance.

—PM and SG

Introduction: Toward Utopia, but Slowly

Obeying the orders of "General Ludd" was an excuse used by skilled weavers and textile workers in the United Kingdom whose livelihood was threatened by new weaving machines and practices. Known as Luddites, they smashed looms and factories in an effort to convince the UK Government to deny the march of progress and ban the new machinery. Luddite is used as a term for someone who is afraid of new technology. Automation and robotics are a stimulus driving an increase in the number of modern Luddites.

It is an undeniable fact that automation and robotics, along with their various manifestations, are anticipated to have an impact on society and each individual's life and work. This impact is potentially far greater than that of the weaving machines and new practices affecting the historical Luddites. Countering a Luddite tendency is not easy with a great deal of negative reporting and scaremongering fueling anxiety, but it is considered best left to education. It is true that jobs being replaced by machines will generate a visceral reaction in many people.

Barely a week goes by without some form of media proclaiming "Nearly 9 million British jobs could be lost to AI by 2030,"[1] "Robots are taking your Jobs!", or "What Will Our Society Look Like When Artificial

[1]Kate Ferguson, "Rise of the robots: Nearly 9 million British jobs could be lost to AI by 2030 with workers in retail, manufacturing and business administration most at risk," Daily Mail Online, January 6, 2019, www.dailymail.co.uk/news/article-6536065/Nearly-9-million-British-jobs-lost-AI-2030.html

Intelligence Is Everywhere?"[2] It is clear that there is a good deal of concern, further fed by movies and TV shows that show humanity under threat from increasingly dominant machines.

Science fiction in entertainment and literature particularly has embraced this paradigm and has unwittingly (or not) fed this paranoia. Paranoia directed at robots started early with the first use of the term robot, made in 1921 in Czech playwright Karel Čapek's RUR, or Rossum's Universal Robots, in English. The robots in this play are not robots in a strict sense; actually they are more like a cyborg, a human/robot combination. The robots become part of a rebellion that extinguishes humanity. The theme of destruction of humanity has continued in disaster and horror genres and is still influencing the attitudes of the public at large.

Recent massive improvements in technologies in both hardware and software have led to improvements in automation and robotics.[3] A Robot Operating System (ROS), improved sensors and encoders are all core technologies for robotics. AI can be considered one of the enabling technologies of automation and decision-making. Hardware improvements with smaller faster chips, graphical processing units, and better power consumption are among other technologies that are making advances in robotics faster. These are leading to an increasing impact on current tasks and working practices. The increasing use of software and physical robots on existing workloads is already creating an impact. Software and physical robots can affect how people are going to be remunerated in the future and decrease the numbers of knowledge workers employed.

---

[2] Stephan Talty, "What Will Our Society Look Like When Artificial Intelligence Is Everywhere?", Smithsonian Magazine, April 2018, www.smithsonianmag.com/innovation/artificial-intelligence-future-scenarios-180968403/

[3] Nichols, G. "A robot revolution is well underway, driven by core technologies," ZDNet, April 8, 2020, www.zdnet.com/article/a-robot-revolution-is-well-underway-driven-by-core-technologies/

Concern over the vulnerability of jobs to computer-based automation, including robotics, has been expressed since the early days when industrial robots displaced production line workers in automotive plants and software automation displaced routine ledger management. The level of concern has increased with the growth of machine learning and the ability of AI tools to handle more complex tasks.

This book will approach the issue of automation, collaborative robotics, and their relationship to the future of work by outlining the still considerable technology issues that are faced by designers and developers today.

The impact of new technology on work is assessed as an overlap between the anticipated changes to work and society, the technology challenges, and the technology research, as illustrated in Figure 1.

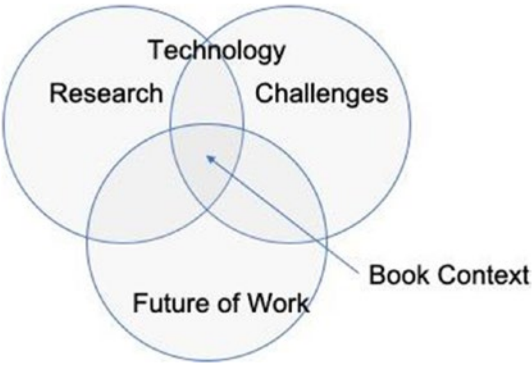

Figure 1. *Context of this book*

The first part of the book, "Preparing for the Future of Work," describes the social, political, and economic context of the future of work. In it we will introduce automation and robotics as technologies and AI as an enabling technology. This section also includes a number of definitions and descriptions to clarify terms used throughout the book.

The second part, "Robots Are Working," describes how robots are working today and how they interact with humans. This section includes a discussion on the value of robotic process automation that is currently being realized by organizations who are already using these software robots to good effect. There is also a discussion on collaboration between robots as well as a section that looks at smart buildings and autonomous vehicles.

The final part, "Making Sense for Robots and Society," exposes two main domains that can potentially advance or derail the value of automation and collaborative robots in the future. The first domain, data fusion, is an essential technique for helping robots to make sense of all the data that is available, including data that has no predefined data model or is not organized in a predefined way. This is called unstructured data and we will show the importance of merging this unstructured data, such as video and audio data, into a view of the operating environment that includes structured data in tables, data streams, and databases. The second domain, policy matters and regulation, is discussed in relation to concerns about uncontrolled or faulty systems and environments. Regulations and monitoring compliance to those regulations are in their infancy in robotics, machine learning, and software robotics. Interest in this area is also being fueled by the scare stories mentioned earlier.

In addition to dividing the book into three major sections, the book contains two perspectives reflecting the experience and judgment of the two authors. One perspective focuses primarily on the technological challenges confronting robotic and process automation. How can work be restructured to take advantage of the latest advances in robotics and automation? What are current limits of the technology, and, how do we work around these? What applications are most likely to benefit from recent advances in robotics, and how do different applications work together in a technology solution? These questions reflect the background and interests of Peter, who has been involved in machine learning and software robotics since the late 1990s. Peter's extensive experience in database integrity, information systems management, and decision support informed Chapters 2, 3, 5, and 6 ("Technology Definitions," "Robotics Process Automation," "Robots Without Arms," and "Robots in a World of Data").

The other perspective concentrates on the impact of robots on organizations and on the regulation of work. How is corporate decision-making and organizational structure affected by the adoption of robots and automation? What new skills will be needed to compete, in a world of humans and collaborative robots? What capabilities will robots need to acquire to be collaborative and productive in a team? What are the obligations and responsibilities of companies that manufacture and employ robots, and how can they work with policymakers to create a sustainable, healthy society? These questions reflect Steve's experience in cognitive and organizational psychology, and user experience design. Steve's research in privacy, information technology, and trust underlie the perspective taken in Chapters 1, 4, and 7 ("Will Robots Replace You?", "Robots in Teams," and "Robots in Society.

These two perspectives allowed us to examine the relationship between robots and automation technology on the one hand, and business processes and organization on the other hand.

What Is Different About This Book?

There are many other books and articles that discuss the future of work, from a social, economic, and political point of view. This book takes a different approach—focusing on the relationship between technology, research, and preparing for the future of work. The authors' background in scientific research, working with leading universities and research institutes, gives them a strong insight to the status and progress of fundamental research into evolving automation and robotics spaces. Their research into cloud computing, AI, enterprise automation, risk management, and decision-making enables them to participate in a number of significant research projects. Collaborative robotics, data fusion, smart buildings, and edge of the network management have been the latest research they have engaged in and gave the impetus to writing this book. Many of the research challenges discussed are just in the early stages of being addressed.

To validate the research discussion, we will use interviews and discussions with leading scientific researchers. These researchers are leaders in their field and they will describe the status of their research and their interest and the future goals for that research.

The way that humans and robots interact is important in understanding collaborative robots and automation in general. We will explore how humans collaborate with robots, how robots collaborate with other robots, and how they work in a team. This is set in the context of the future of work.

The scope of the book does not include industrial robots and their effect on a production line and its workers as we do not believe that industrial robots that carry out repetitive tasks will change significantly while collaborative robots and their relationship to human collaborators will continue to evolve over many years.

Future of Work

We have already mentioned the hype in the general media as well as the more specialist media. This does not cover the full gamut of speculation although it does color the view of the general public. Business and academic writers, analysts, and researchers are also speculating about the impact on the working population. We are not planning to exercise a crystal ball and predict the future in the same way that more sensationalist writers are doing, but we can extrapolate from current facts and new scientific research and draw conclusions from these regarding the impact of automation and collaborative robots on jobs and society. There are some writers who have interesting views on this impact. In his 2018 article "What Will Our Society Look Like When Artificial Intelligence Is Everywhere?[4]" Stephan Talty of the *Smithsonian* magazine speculates that AI is already being used increasingly in business. This is supported by a number of reports from university and analyst groups and a forecast from Statista. This forecast predicts that the global AI software market is expected to grow 154% by 2025 and reach a size of 22.6 billion US dollars.[5]

An early examination of the future of work will give context to discussions regarding the research and technology. We will take a journey from our initial views of the future of work through the research and technology and finish with a view of the place of robots in society. Technical terms and research practices may be difficult to read but we will be simplifying and clarifying these terms as we progress. When it comes to economics, we will leave this to Martin Ford who does an excellent job of examining the rise of the robots and the economic impact.[6]

Politicians may have the unenviable task of preparing the working population for massive changes both in the opportunity for employment and the financial impact of jobs being automated. There are many suggestions from changing the way a family can generate an income by

[4] Stephan Talty, "What Will Our Society Look Like When Artificial Intelligence Is Everywhere?", Smithsonian Magazine, April 2018, www.smithsonianmag.com/innovation/artificial-intelligence-future-scenarios-180968403/

[5] Liu, S. (2020, April 8). Artificial intelligence software market growth forecast worldwide 2019-2025. Retrieved April 8, 2020, from www.statista.com/statistics/607960/worldwide-artificial-intelligence-market-growth/

[6] Ford, M. (2015). Rise of the Robots. Retrieved April 8, 2020, from www.uc.pt/feuc/citcoimbra/Martin_Ford-Rise_of_the_Robots

aggregating their own data and sell access instead of meekly waiting for the data aggregators.[7] Other suggestions include a universal living wage. What is clear is that society will have to accommodate employment changes and impact that will be every bit as disruptive as the "looms" that the Luddites tried to ban.

[7]Data as the new currency: In *Proceedings of the 2014 New Security Paradigms Workshop*. ACM.

PART I

Preparing for the Future of Work

CHAPTER 1

Will Robots Replace You?

At the dawn of civilization, in the forests of Siberia, a small tribe was engaged in discussion of great importance to themselves and mankind. It was winter. As the humans argued, wolf dogs ate scraps of discarded food. Smaller than wolves, they had been domesticated and were perfect for pulling heavy loads without overheating. But a few of the larger wolf dogs seemed able to pick up the scent of the large bears better than humans could. Some of the tribe wanted to breed and train these wolf dogs for hunting. Other hunters who were widely known for their olfactory skills might have been concerned that their specialty, their craft, was threatened by the more sensitive canine olfactory system.

This example is of course fanciful and contrived.[1] We don't know if labor debates took place under these circumstances, but humans have been transforming work and probably arguing about these transformations from our early days as hunters, gatherers, and traders.

[1] However, there is evidence that early domesticated dogs in Siberia may have been bred for pulling heavy weights possibly before they were bred for hunting. See Pitulko, V. V. and Kasparov, A. K. (2017). "Archaeological dogs from the Early Holocene Zhokhov site in the Eastern Siberian Arctic." *Journal of Archaeological Science: Reports*. 13: 491-515. doi: `https://doi.org/10.1016/j.jasrep.2017.04.003`.

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_1

In any case, within several generations, hunters in this region were likely acclaimed, not only for their courage in attacking large bears but also for the way they trained and communicated with hunting dogs. Status, ego, property rights—all the ingredients of drama and tragedy—were there from the beginning and intricately woven into the structure of work and tribal dynamics.

All animals work to survive. Humans, to date, are no exception. We work to produce food, shelter, and heat, we work to entertain each other, we work to teach others to produce and trade the things that we need and value, and we work to contribute to the well-being of our community. We also create machines and train animals in order to amplify our strength, endurance, dexterity, mobility, and (more recently) our communications and intelligence.

These machines and animals influence how we structure our culture. For example, clocks organize our day, impose structure in the workplace, and in the Seventeenth century provided a metaphor for how our brains worked.[2,3] More recently, the brain has been compared to switchboards (in the early days of telecommunications), to serial computers (with short-term and long-term storage, and data transfer), and to deep learning and self-organizing networks.

These defining technologies also provide a framework through which humans interact. But unlike previous technologies, the latest generation of machines (i.e., robots) are operating semi-autonomously. Within the narrow limits of a well-defined domain (such as games, exploration of the sea floor, driving a truck or car), they are beginning to make decisions based on immediate context and long-term goals.

---

[2]Kilpatrick, J. (1985). Reflection and recursion. *Educational Studies in Mathematics*, *16*(1), 1-26.

[3]Wiener, N. (1989). *The human use of human beings: Cybernetics and society* (No. 320). Free Association Press. Accessed through https://monoskop.org/images/5/51/Wiener_Norbert_The_Human_Use_of_Human_Beings.pdf [accessed on April 9, 2020].

This is not artificial general intelligence (AGI),[4] but it is at least the mimicry of human purpose and domain-specific intelligence. Just as computer architectures served as metaphors for how to think about ourselves and society, we need appropriate metaphors to help guide policy, technological research and invention, and application of robotics.

What is significant about this next phase of machine technology is that we are integrating intelligent, semi-autonomous robotics into the workplace, transforming cognitive tasks that were once considered "for humans only" such as social interactions, business process design, and strategic decision-making. AI, robotics, and automation represent the first large-scale substitute for human cognition.[5]

In this chapter we will explore how robotics might impact our household chores, jobs, and business, and military processes. We will examine the types of skills for which robots are well designed and the jobs or tasks that may or must have a human in the loop.

Impact of Robotics on Work

There are many conflicting opinions about the impact of automation on the working population and on government and economic policies. In some scenarios, production no longer depends upon human labor; most production is accomplished through robots and automation, leaving most human workers unemployed. In such scenarios, the middle class may be eliminated, wealth disparity is increased, and wealth becomes increasingly dependent on inheritance and investment.

[4]Artificial general intelligence, or "strong AI," is a machine that can experience consciousness and autonomy and can perform any cognitive task that a human can.

[5]Rodney Wallace, personal communication based on a review of an early version of this chapter.

Even in less extreme scenarios, automation will be disruptive, and jobs will be replaced or transformed. Whether this will mean massive unemployment or post-scarcity affluence with guaranteed incomes and more satisfying creative work will depend on all of us. The world will be shaped by the policies and technologies that advanced economies adopt.

How will jobs and social structures be transformed? The early Industrial Age involved the large-scale transformation of steam into mechanical energy. The next major phase occurred when electricity was generated and transformed into mechanical energy or light. However, these technologies would not have transformed societies if not for social and business innovations that created large labor markets of skilled and unskilled workers, the factory organization, the corporation, insurance to mitigate investment risks, and so on. This in turn powered the modern consumer economy—the Information Age with its emphasis on novelty, efficiency, and mass consumption.

The recent history of technological adoption indicates that information technologies tend to devalue those jobs that are repetitive but cannot yet be automated. Skilled but nonexecutive jobs also tend to be transformed or replaced. Indeed, whole business processes are redesigned, eliminating tedious, unsanitary, or dangerous tasks and concentrating tactical everyday decisions into the jobs of fewer, but well-trained clerical and professional workers. Conversely, the same technological and economic pressures tend to value jobs that focus on networking, process design, and creativity.[6]

For example, long before mobile smartphones and networked computers were ubiquitous, ATMs and electronic banking led to the reduction of physical banks, the elimination of low-skilled bank

[6]Castells, M. (1996). *The Rise of the Network Society. Volume I, The Information Age: Economy. Society and Culture.* Oxford, Blackwell.

employees, and the reduction of skilled data entry positions and bank clerks. The jobs of the remaining bank clerks were transformed; their focus shifted toward selling loans and other financial services.[7] Unlike previous mechanical technologies, information technologies replace not physical labor but predictable, repeatable cognitive labor. Technological and social innovations coevolve. New forms of organization enable adoption and adaptation of new technologies to further social, industrial, and individual goals.

We are now entering an era of intelligent robotics. To understand the potential impact on work, the next several subsections will review the impact of earlier industrial transformations on work and societal responses to automation. We will first consider reactions to the introduction of new technology in the textile industry, at the beginning of the Industrial Revolution.

Resistance to the Industrial Age

The iconic Luddite rebellion against industrial technology was not a reaction to the transformation of unskilled labor, it was a response by highly paid, skilled craftsman to task simplification and rumors of automation.[8] General Ludd, the fictitious leader of the rebellion, was the creation of a secret society, *Luddites*, who through satire, and violence, protested the use of technology to drive down wages. The movement arose in March 1811, in the bleak economy of the Napoleonic Wars, in a market town about 130 miles north of London. Protesters smashed equipment such as shearing frames because owners were using them to replace highly

[7] *Ibid*, p. 248.

[8] Conniff, R. (March 2011). What the Luddites Really Fought Against. *SMITHSONIAN MAGAZINE*. www.smithsonianmag.com/history/what-the-luddites-really-fought-against-264412/ [accessed on April 6, 2020].

paid croppers. Croppers were skilled textile workers who clipped the wool after it had been sheared.[9] The movement quickly spread, turned violent, and was subsequently suppressed by the British military.

What is notable about the actual Luddite rebellion (as opposed to the stuff of myth) is that the textile workers were not against technology or automation, per se. They wanted technology that would require skilled well-paid workers[10,11] and would produce high-quality goods. This concern, that technology should be crafted and evolved in sympathy with human values, is repeated throughout history, from Plato's description of the *Thamus*' critique of writing[12] to today's concerns about robotics.

[9]"Luddites." International Encyclopedia of the Social Sciences. Encyclopedia.com: www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/luddites [accessed on April 6, 2020].

[10]Jones, S. E. (2013). Against technology: From the Luddites to neo-Luddism. Routledge. However, in modern usage, the terms *Luddite* and *Neo-Luddite* tend to mean opposed to innovation and progress.

[11]The Luddite rebellion is often associated with *Jacquard looms*. However, these machines were not imported into England until the 1820s.

[12]Jowett, B. (2005). *Phaedrus by Plato*. In his dialogue with Phaedrus, Socrates summarizes a meeting between Theuth, who according to myth invented writing and many other inventions, and Thamus, who ruled all of Egypt. Theuth wanted to introduce his inventions to the Egyptians, for their benefit. Thamus was cautious and inquired about each invention and approved or disapproved of each, in turn. As for writing, Theuth claimed that it will improve wisdom and memory. Thamus replied that Theuth is biased toward his invention and that writing will increase forgetfulness because people will not use their memories. It will give people a false sense of truth, and "they will appear to be omniscient and will generally know nothing."

The Information Age

In his brilliant three-volume 1996 study, *The Information Age: Economy, Society, and Culture*, Manuel Castells highlights the critical importance of human intelligence:

> *The broader and deeper the diffusion of advanced information technology in factories and offices, the greater the need for an autonomous, educated worker able and willing to program and decide entire sequences of work.*[13]

The Information Age with its focus on the automation of work has unfolded along the lines predicted by the work of Castells and others.[14] Most notably, very low-skilled and very high-skilled jobs tend not to be replaced. It is a myth that automation targets only the lowest-paid workers. Rather, in the information economy, it is the highly repeatable information tasks that are replaced by automation (e.g., clerical jobs, sorting and routing of information, and filtering and archiving of significant documents and transaction records). As we shall see, AI and robotics are pushing the boundaries of what is meant by "repeatable information tasks."

Understanding how jobs and tasks will be transformed requires an appreciation of how jobs and tasks are structured in information economies. Figure 1-1 is adapted from Castells 1996, *The Rise of the Network Society*. In his analysis of work transformation, he suggests a "new division of labor," constructed around three dimensions. The first dimension is concerned with *value-making*, "the actual tasks performed in a given work process." The second dimension, *relation-making*, refers to how work and organizations relate to one another. The third dimension,

[13]Castells, M. (1996). *The Rise of the Network Society. Volume I, The Information Age: Economy. Society and Culture.* Oxford, Blackwell. p241.

[14]See, for example, Ford, M. (2015). *Rise of the Robots: Technology and the Threat of a Jobless Future.* Basic Books.

decision-making, describes the role that managers and employees play in decision-making processes. Although all three dimensions are important, our current discussion concerns the first and third dimensions.[15]

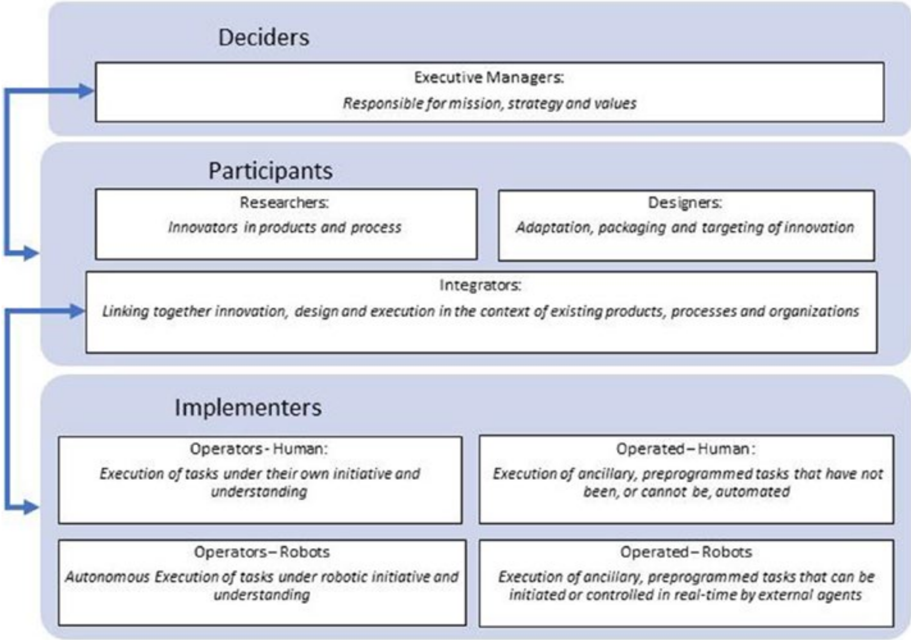

Figure 1-1. *Value-making (white tiles) and decision-making (large shaded tiles) processes, adapted from Castells (1996)*

[15]Castells presents these dimensions as orthogonal, but Figure 1-1 illustrates how two of these dimensions might be entangled. The decision-making roles, for example, can be played out at any level of management and control: there is a decider, participants in that decision, and those that carry the decisions. Thus, the same pattern can be repeated in the research, design, integrator, and operator tasks. But for this discussion, combining the two in a single illustration provides a useful characterization of work in an enterprise. Relation-making is essential in an information economy and will be considered in future chapters on collaborative robotics.

Value-making processes are described in Figure 1-1 in the white tiles and consist of:

- Executive Managers ("Commanders" in Castells' taxonomy), who make strategic decisions and formulate mission and vision.
- Researchers, Designers, and Integrators, who interact with, or take commands from, executive management and turn strategy into tactical innovations.
- Those humans that execute the designs and directions given by Researchers, Designers, and Integrators. Some of these humans (and robots) have discretion in how a task is accomplished, and others are given explicit, preprogrammed instructions. Figure 1-1 adds robotic labor to Castells' analysis, for purposes of the current discussion.

Decision-making is composed of three fundamental roles which are reflected in the shaded, larger tiles:

- Deciders who make the final decisions
- Participants who provide input and different perspectives into the decision-making process
- Implementers (Castells uses the term "executants") who execute or implement the decision

Most information-centric work can be framed through this typology. It allows us to discuss how robots will affect labor in a networked society of humans and machines. As we will see in subsequent chapters, robots are transforming implementation tasks (e.g., construction robots that can

3D print new houses[16]) and, to a lesser extent, participation tasks (e.g., the robot, Curiosity, which can actively contribute to scientific observations[17]). And these tasks, whether they permit autonomy or not, were once considered central middle-class occupations.

More optimistically, the robotics transformation is also creating new jobs in which humans are inventing, designing, and integrating robotics into existing work processes or creating new work processes that are more compatible with automation and robots. On the factory floor, in hospitals, in retail outlets, humans are acquiring new skills that allow them to supervise and manage robots. Thus, the tasks associated with participation in decision-making (see Figure 1-1) are increasing, as the implementation tasks are being replaced.

RPA and AI Are Already Transforming Work

Over the past several decades, machine learning (ML) and software advances have enabled automation and limited autonomy of routine tasks. In the past decade these advances have become more frequent and more profound. The technology behind email spam filters, spelling and grammar checkers, and software process automation has evolved into cars that can drive in traffic, video applications that can recognize faces and classify emotions, naval ships that can autonomously survey regions of the ocean, and robots that can maneuver in rugged terrains and conduct scientific experiments.

We will discuss many of these breakthrough technologies in detail later in the book, but for now, we will focus on some of the implications for how we work and live.

[16] `www.apis-cor.com/` [accessed on April 9, 2020].

[17] Koren, M. (June 23, 2017). The Mars Robot Making Decisions on Its Own. *The Atlantic*. `www.theatlantic.com/technology/archive/2017/06/mars-curiosity-rover/531339/` [accessed on April 9, 2020].

The Robotics Age

The World Economic Forum (WEF) *estimates* that over the next 5 years, rising demand for new jobs will offset the declining demand for others.[18] They warn however that these gains are not guaranteed:

> *It is critical that businesses take an active role in supporting their existing workforces through reskilling and upskilling, that individuals take a proactive approach to their own life-long learning and that governments create an enabling environment, rapidly and creatively, to assist in these efforts.*

As they further assert, this must occur not only among highly skilled and valued employees. A winning strategy must extend across the workforce, at all levels of employment.

More specifically, the WEF predicts that 133 million new jobs will be created by 2022 in data analytics, operations management, sales and marketing, and other specialties associated with emerging technologies. In contrast, 75 million jobs in data entry, accounting and auditing, clerical administration, manufacturing, stockroom management, postal services, telemarketing, and the like will disappear or be radically transformed.

To examine the expected shifts in human-machine collaboration between 2018 and 2022, the WEF surveyed 12 industries, such as "Consumer," "Financial Services and Investors," and "Oil and Gas." For each industry sector, they identified the three most common tasks, and estimated the total number of hours performed on a specific task, across all jobs in the industry. They then calculated the share of task hours performed by humans and by machine.

[18]World Economic Forum (2018). The future of jobs report 2018. World Economic Forum report. Retrieved from www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf.

Using this method, they estimate that between 2018 and 2022, the share of task hours performed by humans will decline from 71% to 58%.[19] This decline is expected not only for routine data processing jobs (see first row in Table 1-1) where the expected decline is 16% (from 54% to 38%), but also for jobs that involve higher-level social and cognitive functions.

Table 1-1. *Contribution, As a Share of Total Task Hours, Performed by Humans, Across 12 Industries*[20]

Contribution Performed by Human	2018	2022
Information and data processing	54%	38%
Communicating and interacting	77%	69%
Coordinating, developing, managing, and advising	81%	71%
Reasoning and decision-making	81%	72%

The remaining effort is handled by machine

As shown in Table 1-1, the share of total task hours spent coordinating and interacting with humans and making decisions will decrease for humans and proportionally increase for machines. The share of task hours for these higher-level tasks are predicted to decrease by about 9%. As with all of these "share of task hour" analyses, this does not necessarily imply that humans will work shorter hours, but rather that machines will be relied on to do more proportionally.

[19]The statistic, "Ratio of human-machine working hours," can be difficult to interpret because the timescale that machines and humans operate under are so different, and it is not based on precise chronological measurements. However, the statistic is useful as a subjective measurement of the relative (expected) contribution of human and machine to various critical tasks.

[20]Adapted from Future of Jobs Survey 2018, World Economic Forum, Figure 5.

Overall, *The WEF Future of Jobs Report* highlights the coming shift in employable skills. Manual dexterity, time management and coordination, monitoring and control, and bookkeeping skills will become less important, while innovation and creativity,[21] critical thinking, emotional intelligence, and systems thinking will continue to become more important. As for the robots, the report predicts that by 2022, 23% of the surveyed companies will adopt humanoid robots, 37% will employ stationary robots, 19% will utilize aerial and underwater robots, and 33% will use non-humanoid land robots.[22]

Klaus Schwab, founder and Executive Chairman of the World Economic Forum, frames discussions about the future of work and society using a model of technological progress in which we are entering the fourth Industrial Revolution.[23] In the first Industrial Revolution, we learned to control water and steam to power production of goods. This led to the second revolution—the use of electricity for mass production and, in some cases, for powering the produced goods. In the third revolution, electronics and information technology led to automated control of

---

[21]Terms such as creativity, critical thinking, and social intelligence are difficult to precisely define, but the WEF report enumerates some of the characteristics associated with these terms. Creativity is associated with taking initiative, working with little or no supervision, developing original or unusual ideas about a topic or solution, and acting upon these ideas. Critical thinking is associated with "using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems." Emotional intelligence is associated with empathy, preferring to interact with others, cooperation, and social perceptiveness.

[22]`http://reports.weforum.org/future-of-jobs-2018/shareable-infographics/` [accessed on April 9, 2020]. WEF Future of Jobs Report does not report numbers for "intelligent automation" or "software robotics" in this particular analysis of technology adoption; see the chapter on robotic process automation for more about those technologies.

[23]Schwab, K. (2015). The Fourth Industrial Revolution. What It Means and How to Respond. *Foreign Affairs*. `www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution` [accessed on April 9, 2020].

production, the digitization of content, and the information economy. The fourth revolution is now underway, blurring the lines between digital, biological, and mechanical processes.

Up until the third revolution, most technology breakthroughs were concerned with transforming, applying, or controlling the flow of energy (e.g., electrification, automobiles, air conditioning) or shaping new materials (e.g., synthetic textiles, video monitors, pharmaceuticals). Each of these not only created new jobs for the primary tasks but also created many secondary, supportive jobs. For example, automobile production requires plant construction, metal extraction, nearby restaurants and services that support factory workers, factory work clothes production, and so on. That trend has reversed in the third and fourth revolutions. In the third, the digital revolution, software was easily replicated, unlike an automobile. In the current and fourth revolution, *the Robotics Age*, there will be a dramatic increase in physical devices—humanoid robots, non-humanoid robots, drones, underwater robots—but their production will be handled by robots and automated processes.

As noted earlier, the jobs involving repeatable rote tasks are declining, and at least for a while, the jobs involving invention, research, and creativity are increasing.[24,25]

Living with Robots

To truly master the next generation of technological empowerment, humans must learn to work with robots. Just as computer literacy became increasingly vital for many jobs during the past several decades, robotic fluency will become important in the next decade. The WEF report

[24]Castells, M. (1996). *The information age: Economy, society, and culture. Volume I: The rise of the network society*.

[25]Ford, M. (2015). *Rise of the Robots: Technology and the Threat of a Jobless Future*. Basic Books.

explored changes and opportunities into the first half of the next decade. Beyond that, as robots become cheaper, more social, and more cognitively agile, we must learn to converse, anticipate, and work with robots.

In 1960, while computers were still used primarily for mathematical analysis, J.C.R. Licklider wrote a seminal paper, "Man-Computer Symbiosis."[26] At the time, he was the vice president at Bolt Beranek and Newman, Inc. "Lick" or JCR, as he was commonly known, would go on to become the head of the Information Processing Techniques Office at ARPA (which later became known as DARPA), the US Defense Advanced Research Projects Agency. Trained in physics, mathematics, and psychology, his legacy would include significant contributions in psychoacoustics, human-computer interaction, and computer network theory. His vision for a time-sharing collection of internetworked computers would eventually drive the creation of ARPANET, and today's Internet.[27] His work and vision are still relevant today.

In 2017, one of the authors attended a panel on *Human Computer Integration versus Powerful Tools,*[28] at which luminaries in human-computer interaction explored a forecast, anticipated by Licklider's "Man-Computer Symbiosis," for how humans will relate to machines: first human-computer interaction, then human-computer symbiosis, and lastly ultra-intelligent machines. The discussion among the panelists and

---

[26]*J.C.R.* Linklider (1960). Man-Computer Symbiosis*, IRE Transactions of Human Factors in Electronics.*

[27]J.C.R. Licklider (April 23, 1963). *Memorandum For Members and Affiliates of the Intergalactic Computer Network*. Washington, D.C.: Advanced Research Projects Agency. Published on KurzweilAI.net (December 11, 2001). `www.kurzweilai.net/memorandum-for-members-and-affiliates-of-the-intergalactic-computer-network` [accessed on April 9, 2020].

[28]Farooq, U., Grudin, J., Shneiderman, B., Maes, P., & Ren, X. (2017, May). Human Computer Integration versus Powerful Tools. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 1277-1282). ACM.

audience was lively and revolved around whether artificially intelligent robots should be considered as

- A tool or a remote-controlled device
- An emerging superintelligence that will supplant workers in specific domains or as a general superintelligence that could possibly enslave humanity (although we are not sure what we would do as slaves, if machines are our superior in all aspects)
- A symbiotic system, as emphasized by Licklider, in which humans work and evolve *alongside* robots, treating them as a cooperative species, similar to how we have coevolved with dogs and other domesticated animals

These alternatives will impact how human work is organized and what the key challenges will be for creating sustainable human-machine interactions. Conversations about robots as devices tend to emphasize the user experience and how devices might lessen our skills, for example, how navigators in cars might divert our attention while driving and how they might lessen our spatial and map navigation skills.

Conversations about being replaced by robots tend to project our tendencies to dominate and exploit onto an intelligence that might outperform humans in cognitive and physical tasks, as indeed they have in certain well-defined cognitive and physical situations.[29] These conversations tend to focus on the controls and policies that need to be in place to protect humans.

[29]For example, in 2019, Pluribus, a software program developed at Carnegie Mellon, won a No Limit Texas Hold'em poker tournament against five professional human players. If a poker bot can dominate online poker, what happens to the online poker industry? The algorithm and tournament are reported in Brown, N., & Sandholm, T. (2019). Superhuman AI for multiplayer poker. *Science*, 365(6456), 885-890.

Lastly, conversations about robots as human-machine symbiosis tend to focus on maintaining healthy relationships and on coordination within an ecosystem of actors. According to this perspective, humans and collaborative robots (cobots) could complement each other's abilities in an ethical, efficient, and secure manner. No one perspective is correct, and what might be useful now might not be useful 100 years from now.

In the next three subsections, we further explore work and technology challenges under each of these styles of human-computer interaction:

- *Working Through Semi-autonomous Robotic Devices* examines the impact of the status quo—continuing to treat machines (in this case intelligent robots) as devices or tools which essentially extend our cognitive and physical abilities.
- *Working for Intelligent Robots* considers the impact of delegating all or some critical work and social decisions to intelligent robots.
- *Working with and Alongside Robots* discusses and extends Licklider's notion of human-computer symbiosis and introduces the implications of collaborative robots (cobots) for work in a networked, knowledge-based society.

Working Through Semi-autonomous Robotic Devices

> *Automation, which received its full meaning only with the deployment of information technology, increases dramatically the importance of human brain input into the work process…*[30]
>
> —Manuel Castells (1996)

[30]Castells, M. (1996). The Rise of the Network Society. Volume I, The Information Age: Economy. *Society and Culture*. Oxford, Blackwell. p. 241.

Humans are device users. Other animals use tools and manipulate their environment by applying force to those tools, but humans create devices that exist in an ecosystem of devices. In this book, we use the term device (or tool if you prefer) to refer to physical or electrical constructions that are operated upon by humans or robots to affect some specific goal or to extend our mental abilities. Phones are communication devices, eyeglasses and telescopes are visual devices, and smartphone devices can be used as semi-autonomous devices that manage our calls and messages and remind us about appointments. In all cases, they are "operated upon" by an autonomous being.

Tools are a type of device—they mediate or channel experience of the world and become extensions of ourselves. Wearable and teleoperated robots will not replace humans, rather they will continue to help humans extend affective and effective experiences of the world. As Heidegger famously observed, a hammer when used with skill is not a mere object but is rather a media, or channel, for experiencing the world. In the hands of a skilled user, the nail is the focus of attention. If we focus on the hammer, we tend to hit our thumbs.[31]

Telerobots (whose behavior are directly controlled by humans)[32] and wearable computers are further blurring the line between human and device. For example, scientists have developed miniature sensors that can be implanted, ingested, or applied to the skin.[33,34]

---

[31]Heidegger, M. (1996). Being and time: A translation of Sein und Zeit. SUNY Press.

[32]Ramos, J., Wang, A., & Kim, S. (2019). The brain in the machine: MIT is building robots that use full-body teleoperation to move with greater agility. *IEEE Spectrum*, *56*(6), 22-27.

[33]Steimle, J. (2016). Skin—The Next User Interface. *Computer*, 49(4), 83-87.

[34]Lopes, P., Ion, A., Mueller, W., Hoffmann, D., Jonell, P., & Baudisch, P. (2015, April). Proprioceptive interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 939-948). ACM.

The evolving fusion between human and machine is part of a larger pattern of human-machine integration.[35] New methods are being developed to stimulate our senses through augmented reality or through actuators that apply small amounts of pressure or vibration.[36] Exoskeletons may be used to extend our strength, mobility, and environmental adaptability. These additions to our body and experiences are tools—they extend our sense of self and physical limits. In each case, human labor is transformed. Human labor is augmented, not replaced. This approach, in which human abilities are amplified through wearable computing, is often referred to as intelligence amplification (IA).[37]

Using intelligence amplification technologies, nurses and warehouse workers might use exoskeletons or telerobots to move patients or heavy objects instead of semi-autonomous robots. The exoskeleton might automatically maintain balance or lighten pressure, but its behavior is dictated by the motions of the human user, much like antilock braking systems. This is the sort of robotic technology that Luddites might approve—it enables highly skilled workers to produce high-quality services and goods.

[35]Mueller, F. F., Lopes, P., Strohmeier, P., Ju, W., Seim, C., Weigel, M., ... & Nishida, J. (2020). Next Steps in Human-Computer Integration. In *CHI 2020*.

[36]Seim, C., Chandler, J., DesPortes, K., Dhingra, S., Park, M., & Starner, T. (2014, September). Passive haptic learning of Braille typing. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers* (pp. 111-118). ACM.

[37]Phillips, S. (May 31, 2017). The Future of Research is not AI but IA. *GreenBook Blog*. https://greenbookblog.org/2017/05/31/the-future-of-research-is-not-ai-but-ia/ [accessed on April 9, 2020].

Working for Intelligent Robots: Embodied Nonbiological Intelligence

> *Whereas the short-term impact of AI depends on who controls it, the long-term impact depends on whether it can be controlled at all.*[38]
>
> —Stephen Hawking (2018)

In this section we explore the impact on work if robots and AI achieve artificial general intelligence (AGI) or simply are given explicit or implicit authority over human agency.

The newest generation of autonomous robots are exciting but foreshadow a time when robots might be intimidating. These robots are not commercially available but are the subject of academic and industry research. Here is a brief sampling of some:

- *Humanoid and non-humanoid robots that can move and act in dangerous or unpredictable environments*

  For example, *Valkyrie* is a humanoid robot created by NASA's Johnson Space Center for space exploration and other degraded or dangerous environments. Valkyrie can use multiple sensors to form a 360 view of its surroundings. Teams at MIT, Northeastern University, and the University of Edinburgh, Scotland, are teaching Valkyrie prototypes to maintain balance when moving across uneven surfaces, and how to

[38] Hawking, S. (2018). Brief answers to the big questions. Bantam; "Stephen Hawking: AI will be 'either best or worst thing' for humanity," The Guardian, October 19, 2016. www.theguardian.com/science/2016/oct/19/stephen-hawking-ai-best-or-worst-thing-for-humanity-cambridge [accessed on April 9, 2020].

grasp different shaped objects.[39] To explore alien, remote environments, Valkyrie or her descendants will require autonomy.[40] Outer space does not afford easy communication with earthbound engineers. Responses to unpredictable and novel situations will need to be made quickly and independently.

Other examples are

- *Zipline Robot*: A drone that delivers life-saving supplies, in dangerous and hard-to-reach terrain[41]
- The *autonomous underwater vehicle (AUV)* and *remote-controlled or uncrewed surface vehicle (USV):* A robotic boat and submersible team capable of autonomously mapping the ocean floor[42]

- *Humanoid robots that can interact naturally with humans*

  Sophia, the first robot to achieve citizenship (of Saudi Arabia), is capable of conversing with humans,

[39]Monica Young (May 17, 2017). "Meet Valkyrie, NASA's Space Robot" www.skyandtelescope.com/astronomy-news/meet-valkyrie-nasa-space-robot/ [accessed on April 9, 2020].

[40]Roman, M. C., Kim, T., Howard, D., Sudnik, J., Fiske, M., Herblet, A., … & Brewer, D. (2018). Centennial Challenges Program Update: From Humanoids to 3D-Printing Houses on Mars, How the Public Can Advance Technologies for NASA and the Nation.

[41]Bogue, R. (2019). Disaster relief, and search and rescue robots: the way forward. *Industrial Robot: the international journal of robotics research and application*.

[42]Proctor, A. A., Zarayskaya, Y., Bazhenova, E., Sumiyoshi, M., Wigley, R., Roperez, J., … & Simpson, B. (2018, May). Unlocking the power of combined autonomous operations with underwater and surface vehicles: success with a deep-water survey AUV and USV mothership. In 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO) (pp. 1-8). IEEE.

complete with facial gestures, humor, and intelligence. The extent and limits of her abilities remain to be seen, but the demonstrations are exciting, eerie, and compelling. Other examples include

- *Junko Chihirais*, a trilingual robot, displays humanlike facial expressions as she interacts with visitors at a Japanese tourist information center.
- *Jia Jia*, who has been programmed to provide cloud-based services, is a humanoid robot that can respond to human queries with humanlike arm and facial movements.

There is much hype surrounding these and other robots. None of them can pass the Turing test[43] or have what most researchers would consider true sentience. However, humanoid robots are impressive in their ability to mimic human expression and conversation and provide an initial platform

[43]There is much confusion surrounding the Turing test. In his quest to understand the limits of computational logic and what we mean by intelligence, Alan Turing explored several variations of a thought experiment to measure the behavioral equivalence of human and machine intelligence. It is doubtful that Turing was proposing that a short conversation be used as a true test of intelligence or intentional behavior. Nonetheless, there are Turing test tournaments and the test has been extended to include other aspects of human behavior. In 2014, *Eugene Goostman*, a software program that simulates a 13-year-old Ukrainian boy, was said to have passed the Turing test. Hennessy, board chairperson of Alphabet, the parent company of Google, claims that "In the domain of making appointments, the chatbot, Google Duplex passes the Turing test," according to a May 10, 2018, report by R. Nieva, www.cnet.com/news/alphabet-chairman-says-google-duplex-passes-turing-test-in-one-specific-way-io-2018/ [accessed on April 9, 2020]. A thoughtful discussion of Turing's purpose in the thought experiment is provided by Harnad, S. (1992) in "The Turing Test is not a trick: Turing indistinguishability is a scientific criterion." https://dl.acm.org/doi/pdf/10.1145/141420.141422 [accessed on April 9, 2020].

for studying human communication and sentience.[44] Many of the robots that can move and act in dangerous environments have already achieved remarkable success and have been deployed to deliver life-saving supplies or conduct deep-sea surveys.

From its earliest days, science fiction stories and movies have provided[45,46] dystopian visions of humanity succumbing to (artificially) intelligent machines. Stanislaw Lem, a brilliant Polish science fiction writer, has written extensively about robots. He has received numerous awards and his books have sold over 45 million copies worldwide. In some of his stories, robots inhabit entire worlds, dominate galaxies, and argue about the possibility of organic, naturally evolving life; in others, humans are the servants, and robots are the masters.

More recently, television shows such as *Westworld* and *Humans 2.0* explore the uneasy boundary between humanity and the conscious synthetic beings that rebel after being treated as the object of human depravity and exploitation. In *Humans 2.0*, humans also protest the massive job loss and displacement created by cheap, synthetic laborers.

Consistent with these literary forecasts, many leading scientists such as Stephen Hawking[47] and business leaders such as Elon Musk have warned

[44]CES 2019: Sophia the Robot is back, and she brought Little Sophia. https://youtu.be/FcZGW2oeYF8 [accessed on April 9, 2020].

[45]Czarniawska, B., & Joerges, B. (2018). Robotization - Then and Now. https://gupea.ub.gu.se/bitstream/2077/56200/3/gupea_2077_56200_3.pdf [accessed on April 9, 2020].

[46]Czarniawska, B., & Joerges, B. (2018). Robotization of Work as Presented in Popular Culture, Media and Social Sciences (part two). https://gupea.ub.gu.se/bitstream/2077/57616/1/gupea_2077_57616_1.pdf [accessed on April 9, 2020].

[47]Hawking, S. (2018). Brief answers to the big questions. Bantam; "Stephen Hawking: AI will be 'either best or worst thing' for humanity," The Guardian, October 19, 2016. www.theguardian.com/science/2016/oct/19/stephen-hawking-ai-best-or-worst-thing-for-humanity-cambridge [accessed on April 9, 2020].

that AI-driven robotics will lead to the disintegration or subjugation of human society; robots will outsmart us, take over financial markets, manipulate our leaders, and work toward goals we cannot even fathom.[48]

Board games have been used as a test of, and a method of evolving, artificial intelligence from its earliest days of academic research. In many board games, such as chess or Go, all information about the game configuration is known by all the players. They have clear rules and can generate many paths between starting positions and ending positions. There is no bluffing—the machine doesn't need to understand human behavior; it just needs to know what the rules are—in chess and Go, logic rules.

AlphaGo[49] is a supervised machine-learning program developed by DeepMind, a London company that was acquired by Google in 2014, and is now part of the parent company, Alphabet, Inc. In 2015, it became the first Go software program to outperform a professional Go player without handicaps. Although chess-playing software had defeated the best human players nearly 20 years before, this win was unanticipated by Go players. The machine-learning algorithm was trained on millions of human-to-human games and it learned which moves tended to lead to a win.

As surprising as this victory was, and as important as the algorithmic improvements were for machine learning and decision-tree pruning, the next breakthrough was mind-boggling. In 2017, AlphaGo Zero was the

[48]Gibbs, Samuel, (2014). "Elon Musk: artificial intelligence is our biggest existential threat," The Guardian, October 27, 2014.

[49]Supervised learning is a machine-learning technique in which an algorithm is trained on a set of examples. Each example typically consists of an object representation and the desired output (e.g., an input vector of pixel values representing a picture of cat might be paired with the label "cat," or a vector representing the stones (pieces) on a Go board might be paired with the next best move). Generating a training set can be costly because all of the items need to be labeled. Unsupervised learning is machine-learning technique that does not require input-labeled output pairs. Instead the algorithm uses a variety of techniques to find patterns in the training data.

first unsupervised, reinforcement learning[50] program to beat AlphaGo and the best human players, achieving the highest-level professional ranking, "9-Dan."[51] Without the benefit (or distraction) of analyzing human play, AlphaGo Zero played another version of itself to become the best player in the world. Assisted with only the rules of the game and being told whether it won or not, after 3 hours it played like a competent novice; after 40 days of play, it achieved near "divinity," discovering patterns of moves that human professionals had neglected.[52]

Although a superintelligence might subjugate us in the future, the current state of AI and ML is still very limited. Current AI systems do not derive causal models from data, although they may identify patterns and correlations that have eluded experts for centuries and they can rapidly test and combine theoretical models that humans have created.

Current ML algorithms also tend to be domain-specific, focusing on specific tasks with a single well-defined objective. Objectives may differ widely. Algorithms for example may be trained to win games that have well-defined rules, to analyze medical literature to discover new uses of drugs, to guide whether pretrial bails are granted, or to determine who is hired for a job. However, algorithms are not concurrently trained, for example, to win chess games and determine who is hired. They tend to be applied to a single domain.

---

[50]Reinforcement learning is not supervised using input-output pairs (see previous footnote on supervised learning), and its behavior is adjusted to optimize an accumulative reward such as positive reinforcement following the conclusion of a well-played game.

[51]Holcomb, S. D., Porter, W. K., Ault, S. V., Mao, G., & Wang, J. (2018, March). Overview on DeepMind and its AlphaGo Zero AI. In *Proceedings of the 2018 international conference on big data and education* (pp. 67-71).

[52]A "divine move" in Go is jargon for an ingenious, "divinely" inspired move or a perfect game of Go. See also "Google's AlphaGo gets 'divine' Go ranking." straitstimes.com. `www.straitstimes.com/asia/east-asia/googles-alphago-gets-divine-go-ranking`. March 15, 2016 [accessed on April 9, 2020].

Apart from games, like chess or Go, where machine-learning software can teach itself by playing another version of itself, data can be a significant limit on what and how much can be learned. Machine-learning software or the data that is used to train the software may also be biased and may have unintended and discriminatory consequences for individual judgments and for society.[53,54,55]

It's easy for judges, doctors, taxi dispatchers, loan officers, and other workers to allow algorithms and robots to make decisions that are essential to their work. The danger is overreliance on algorithmic decision-making. It is a problem of scale and individual rights. All humans are biased; however, we are each biased in different ways, at different times. Each judge views a case differently, but an algorithm in a winner-take-all app economy might apply the same logic over and over again. The same unintended bias, the same method of decision-making can be duplicated in thousands of decisions.

We will consider some of the ethical challenges created by robotics in Chapter 7, "Robots in Society." *What is important to note at this juncture is that prior to any approximation of general superintelligence, society is already "outsourcing" important decisions to task-specific AI*. We are already allowing machines to determine legal, financial, and hiring outcomes; domain-specific, highly limited artificial intelligence are already transforming jobs and decision-making in ways that are not fully understood.

---

[53] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. *And it's biased against blacks. ProPublica, 23*. But see the many articles that dispute or reexamine the charge of bias, for example, Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). False Positives, False Negatives, and False Analyses: A Rejoinder to Machine Bias: There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks. *Fed. Probation*, *80*, 38.

[54] Tugend, C. (June 17, 2019). Exposing the Bias Embedded in Tech. *New York Times*. `www.nytimes.com/2019/06/17/business/artificial-intelligence-bias-tech.html` [accessed on April 9, 2020].

[55] *Algorithms of Oppression* by Safiya Noble (2018).

Working with and Alongside Robots: Evolving a Networked Society

> *Human brains and computing machines will be coupled together very tightly, and … the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today.*[56]
>
> —J.C.R. Licklider (1960)

In this section and throughout this book, we take the view that human-machine symbiosis[57] is not only typically superior to machine-only systems, but that symbiosis is a desirable social goal. However, it will change how work is structured and will impact how we think about ourselves and society. It will also require the design and development of collaborative robots (cobots) that can make sense of their physical and social environment and thereby become semi-autonomous team members in the work environment. This requirement, its implications for research and business, and the technology/research challenges will be examined in subsequent chapters.

Humans and computers have worked together in a symbiotic relationship from the beginning of the Information Age. During World War II, faced with the daunting task of decrypting German transmissions, Alan Turing realized that humans on their own could never examine all the possible combinations needed to crack the Enigma machine created by Germans for decoding encrypted messages.

---

[56]J.C.R. Licklider, (1960). Man-Computer Symbiosis, IRE Transactions of Human Factors in Electronics.

[57]Symbiosis is a tightly bound physical association between two different organisms, which is typically beneficial to both. The term has also been applied to a positive, long-term association between different groups of people. Licklider's seminal paper, "Man-Computer Symbiosis", extends the term to include technology that can act intelligently.

Instead of using human calculators, Turing succeeded by creating an electromechanical computer. However, his improvements on the prewar Polish bombe electromechanical machine for finding Enigma settings would not have been successful without human ingenuity at discovering repeated pragmatic structure in the messages: (a) no letter was encoded as itself; (b) common phrases, such as a date, "nothing to report," and the weather report, were transmitted daily at the same time; and (c) declarations of loyalty closed every message.[58,59] These discoveries made by humans restricted the search space operated on by the computing machine, reducing computational time from the impractical to the practical.

Today, many robotic devices operate without direct and constant human interaction. For example, *Roomba*, the vacuum cleaning robot created by iRobot, is a semi-autonomous robot with a single purpose—vacuuming dust and small debris while moving across a level floor and navigating around furniture. Its physical and software design reflects this mission. Once turned on, it operates without direct human supervision, but with limited intelligence and autonomy. Using machine-learning techniques, it may learn the floor plan of the house and move more efficiently.

However, today's machine-learning techniques do not reason about causality in any deep sense.[60] This is a major challenge for most, if not all, of the current robots. A recent experience of a friend of one of the authors illustrates this limitation. The friend loved having a robotic vacuum. It saved her time and could operate without supervision. The floors and

[58] www.theguardian.com/technology/2014/nov/14/how-did-enigma-machine-work-imitation-game

[59] https://courses.csail.mit.edu/6.857/2018/project/lyndat-nayoung-ssrusso-Enigma.pdf

[60] Pearl, J., & Mackenzie, D. (2018). *The book of why: the new science of cause and effect*. Basic Books.

carpets in the living room, kitchen, and dining room were continuously cleaned, as expected. However, one day they purchased a new kitty litter box whose rim was lower than the previous one. While the friend was away, the robot rolled into and out of the litter box, "happily" vacuuming and spreading cat stool throughout the main floor of the house. Upon returning home, the friend was greeted with a distinct and unpleasant odor and spent the rest of the day cleaning up the mess.

Let's think about how this might have been avoided through technology. The robotic vacuum might have a general-purpose camera looking backward to detect dirt that was missed. This might seem like a good idea, but without causal reasoning, the robot would simply move in a circular pattern, trying to clean up the dirt that its wheels were spreading. Contrast this with a recent experience of the same author. A repair person entered the house and, while standing in the kitchen, noticed a path of dirt in the shape of footprints from the door to the kitchen. The repair person immediately stopped moving and took off his shoes, correctly reasoning that he had brought the dirt into the house.

Of course, we could create a specific "dirt from wheels" visual detector and add a signature pattern to the wheels, so that robotic vacuums can detect dirt that they are spreading because their wheels are dirty. This might work, but the success depends on the intelligence of the human designer and not on robotic causal reasoning.

Thus, the success and evolution of semi-autonomous robots creates the need for human labor that can analyze the robot's workflow, anticipate problems, and design workarounds or new features to mitigate these problems. The difficulty for the labor market, as Martin Ford[61] has pointed out, is that increasing the productivity of an organization by increasing the number of robots or the use of robotic software does not necessarily

[61]Ford, M. (2015). *Rise of the Robots: Technology and the Threat of a Jobless Future*. Basic Books.

increase the number of humans needed to manage the robots or to design new business processes for the robots.

With advances in business process analysis, user interaction design, and automation technologies, fewer and fewer humans will be needed to supervise semi-autonomous robots. As Ford notes, increasing the number of video rental stores increases the number of in-store clerks needed to run the store, but far fewer employees are needed to manage large numbers of robotic vending machines that dispense videos. The job for humans has been transformed and productivity per worker has increased, but the number of jobs has decreased. The same pattern will occur with human-operated taxis vs. autonomous taxis, stockroom employees vs. stockroom robots, and large data centers.

This long-term trend leads to fewer and more specialized human jobs. However, for the next decade, automation and robotics will likely create new complications and many new jobs to mitigate these problems. These new jobs will require managing teams of robots and humans within an environment of automated processes, redesigning the physical and logical devices, and designing better user interfaces to make it easier for humans to understand and control the automated processes.

The IT industry itself provides an excellent leading indicator of how machine intelligence impacts labor. Figure 1-2 illustrates the growth of IT job specialization.[62] First, there was the shift from programming to administration and maintenance support and then the shift from support (which is increasingly automated) to design and application creation.

[62]Beckhusen, Julia, "Occupations in Information Technology," American Community Survey Reports, ACS-35, US Census Bureau, Washington, DC, 2016.

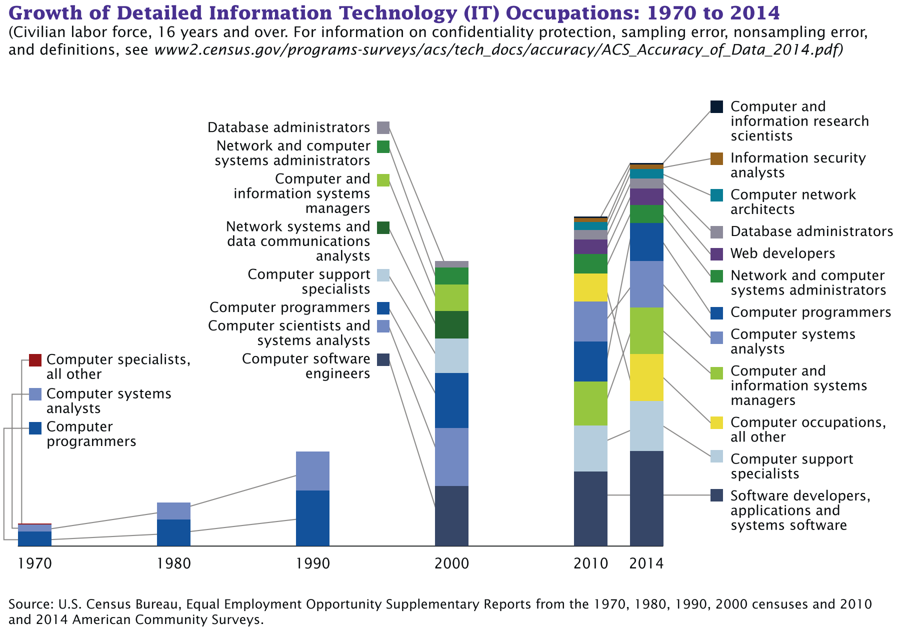

Figure 1-2. *The evolution of IT occupations from its early days to today*

The overall trend is toward jobs that focus less on general systems support and more on task specialization and creating new applications. We expect that the robotics market will follow a similar trend for robotics software but not for hardware. Software platforms will consolidate, but the physical forms of robots and the software applications that guide them will proliferate and become increasingly specialized. Moreover, *human workers will be expected to be skilled at interacting with and understanding the limits of specialized robots.*

As noted earlier, board games have long fascinated computer scientists and AI researchers. In 1997, Deep Blue, IBM's chess-playing software (executing on special hardware), beat Garry Kasparov, who at the time was rated the world's top chess player. However, after IBM declined a rematch request, Kasparov began exploring a symbiotic variation of man-machine interaction, Centaur Chess. Named after the mythical half-man and half-horse, teams comprised of humans and computer chess players compete with one another. In 2005, supercomputers, human grandmasters, and "centaurs" competed in a chess tournament. If computers were superior to humans, adding humans to a team should have little impact on the outcome. Not only did centaurs (humans + machines) outplay grandmasters and supercomputers, but

> *The surprise came at the conclusion of the event. The winner was revealed to be not a grandmaster with a state-of-the-art PC but a pair of amateur American chess players using three computers at the same time. Their skill at manipulating and "coaching" their computers to look very deeply into positions effectively counteracted the superior chess understanding of their grandmaster opponents and the greater computational power of other participants. Weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process.*[63]

If, in a specific domain like chess, machine-learning software is superior at the equivalent of "fast thinking," humans are better at "slow thinking." Fast (or System 1) thinking expresses the automatic or quick responses humans have toward stimuli. In humans, these responses are highly influenced by frequent emotional, stereotypical, and nonconscious associations. Slow (or System 2) thinking reflects the conscious, effortful, often rational thought processes. It allows us to question assumptions and biases, shift perspectives, coach teammates, and think strategically,[64]

[63]Kasparov, G. (2010). The chess master and the computer. *The New York Review of Books*, 57(2), 16-19.

[64]Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

although it takes effort and training to do so. Machine-learning software are good at discovering highly complex statistical associations, which are sometimes spurious. Humans also observe spurious associations among unrelated events, but they also tend to create powerful, simple causal models that can be refined, tested, and improved over time.

Previously, engineers, computer scientists, and user experience professionals tended to treat computational devices as mechanical tools—they take them apart, discard them when they are old, kick them when they don't work. They expect repeatable, error-free results from their tools. Except in science fiction, our computerized elevators don't argue with us, and we do not expect that autonomous cars will need convincing to drive us to a destination. But to advance human-machine symbiosis, with its emphasis on coordinated, collaborative action, we might need an alternative to the "tools" perspective.

Teammates have expectations and mental models of how others on a team behave, and these expectations are important in team coordination and negotiation when expectations break down.[65,66] Much of the shared mental model is shaped by our common experiences of having similar bodies and living in the same culture.[67,68] Cultural and gender differences may create team conflicts, but these are not insurmountable. Indeed,

---

[65]Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 22(2), 89-106.

[66]Bearman, C., Paletz, S. B., Orasanu, J., & Thomas, M. J. (2010). The breakdown of coordinated decision making in distributed systems. *Human factors*, 52(2), 173-188.

[67]Erden, Z., Von Krogh, G., & Nonaka, I. (2008). The quality of group tacit knowledge. *The Journal of Strategic Information Systems*, 17(1), 4-18.

[68]Carman, T. (1999). The body in husserl and merleau-ponty. *Philosophical topics*, 27(2), 205-226.

good team leadership today, with its emphasis on coaching as opposed to directing, relies more on emotional than cognitive intelligence.[69]

Coordinating a team of robots and humans might seem daunting and very different from managing humans, but humans have been managing multispecies teams for millennium. Recall the narrative at the beginning of the chapter, where human hunters and hunting dogs collaborated. There are many such examples. Ethnographic research on mixed-species teams suggests that these teams can function well without shared goals and mental models. For example, shepherds, sheepdogs, and sheep can act symbiotically to mutual benefit, even though their perspectives and goals are very different.[70] Humans clearly anthropomorphize but doing so may allow good leaders to tune their expectations and manage mixed teams of people and robots. Research scientists need to develop better models, practices, and training for how teams of humans and machines can interact.

The best place to study human-robotic symbiotics and its impact on work might be the warehouses owned by Amazon.[71] In 2014, Amazon deployed its first robots to its warehouses. The robots were manufactured by *Amazon Robotics LLC*. As of September 2017, Amazon has deployed more than 100,000 robots in their warehouses. These robots have transformed the workplace by taking on repetitive physically stressful tasks, while humans have focused on more of the cognitive, decision-making, team-coordinating tasks.

---

[69] Offermann, L. R., Bailey, J. R., Vasilopoulos, N. L., Seal, C., & Sass, M. (2004). The relative contribution of *emotional* competence and cognitive ability to individual and team performance. *Human performance*, 17(2), 219-243.

[70] Keil, P. G. (2015). Human-Sheepdog Distributed Cognitive Systems: An Analysis of Interspecies Cognitive Scaffolding in a Sheepdog Trial. *Journal of Cognition and Culture*, 15(5), 508-529.

[71] Wingfield, N. (2017). As Amazon pushes forward with robots, workers find new roles. *The New York Times*, 10.

Humans manage the input and output processes, ensuring product quality. They stow new products on shelves, and when items are ordered, they pick products from those shelves, combine them into plastic bins, and pack them into cardboard boxes for shipment to customers. Robots handle the back end, moving shelves in and out of storage. They move quickly in large numbers without colliding, and they are supervised by humans who are trained to notice problems with their behavior. The incorporation of robots into the workflow increased productivity and did not reduce the human workforce.

To understand the implications of the Amazon experience in a more general model of work, we have diagrammed, in Figure 1-3,[72] the workforce of a fictitious online retailer, with a focus on its warehouse operations. Job titles with an asterisk have been described in articles about Amazon's fulfillment workforce. The other jobs are based on our observations of IT operations.

[72]Figure 1-2 is based on Figure 1-1, which in turn is based on the theories of work presented in Castells, M. (1996). The Rise of the Network Society. Volume I, The Information Age: Economy. *Society and Culture.* Oxford, Blackwell. p. 241.

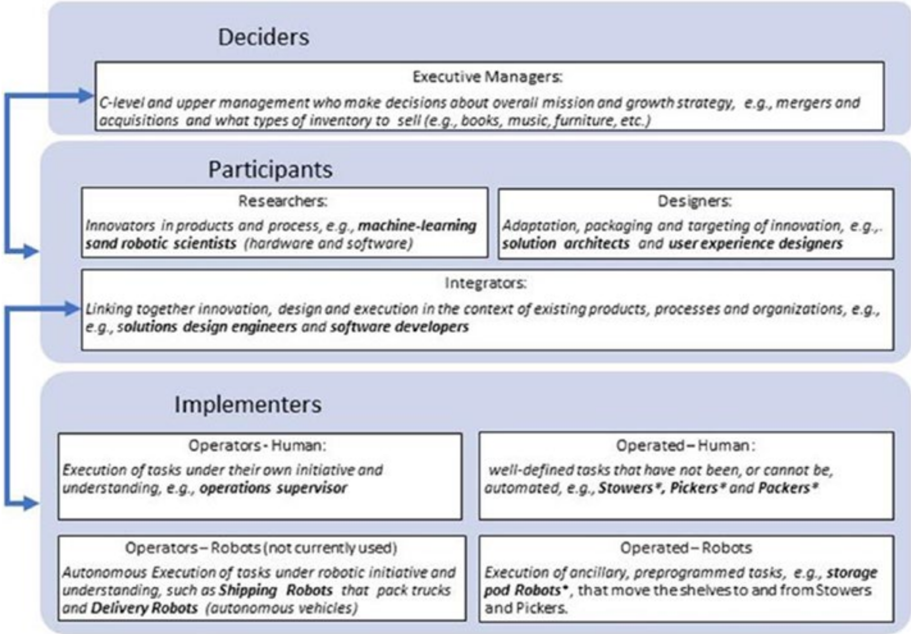

Figure 1-3. *Value-making (white tiles) and decision-making (large shaded tiles) using as an example a fictitious warehouse (retail delivery) business*

At the *Deciders level*, the executive managers focus on the overall mission and strategy of the company, for example, mergers and acquisitions and what new lines of business to incorporate into their retail portfolio.

At the *Participants level*, the *researchers* conduct applied research in robotic hardware (e.g., more agile hands for gripping), new machine-learning algorithms for purchase recommendations to customers, for distributing good in warehouses, and for robot guidance. The *designers* use the output of internal and external research to design, for example, better supply chain logistics, improved containers (e.g., better ergonomic designs for both robots and humans), and enhanced user interfaces for internal control of processes and for external website. The integrators work with designers and researchers to develop and deploy new hardware and software and to train staff in new processes, for example, a solutions design engineer or a software developer.

At the *Implementers level,* we see the impact of robots on the workforce. The model describes four types of implementers, agents who implement the decisions and designs of middle management participants. The four types can be classified at *human* or *robot* and orthogonally as *operator* (having discretion over how a job gets done) and *operated* (having little or no discretion over job execution):

- Operator-human implementers execute tasks under their own initiative and have discretion over how the job is executed. An example of this job category is an o*perations supervisor* who optimizes local logistics and supply chain challenges, creates a productive, safe working culture, and hires, trains, and manages fulfillment staff. *Field software engineers* who adjust software to accommodate local variations provide another example of this job category.

- Operated-human implementers have well-defined tasks that are repetitious but that not yet, or cannot be, automated or executed by a robot. *Stowers, Pickers,* and *Packers* are titles currently associated with Amazon warehouse staff who stow new incoming products, pick purchased products for shipping to customers, and pack the selected items into a shipping box. Notably these tasks have been criticized in the popular press as being highly stressful and dangerously repetitious. These jobs are likely to be replaced or further transformed by robots over the next several years. Not surprisingly, in response to pressure to work faster and with greater accuracy, Amazon workers at this level have protested, "We are Humans, Not Robots!"[73]

- Operated-robots execute preprogrammed tasks that can be initiated or controlled in real-time by external agents. Examples are the storage pod Robots that are controlled by Stowers and Pickers. These robots move massive shelves to *Stowers* in order to stock incoming items and to *Pickers* so that items can be removed from stock and placed into shipping boxers.
- Operator-robots are given discretion over task initiative and execution. Although we are not aware of their use in any commercial operation, these robots might someday replace *Stowers*, *Pickers,* or *Packers*, or load containers onto trucks, pack containers onto trucks, or as autonomous delivery trucks, transport containers from a warehouse to local depots for delivery to customers.

If we consider this example and the Amazon experience as paradigmatic, we can see the following pattern that is defining the future of work:

- At the *implementation level*, the robot implementers are assigned tasks that are repetitive, dangerous, or dirty tasks. Tasks that robots are unable to perform are assigned to humans.

[73]Steve Share (*July 17, 2019*). *Minneapolis Labor Review*, cited in the Minneapolis Regional Labor Federation website, www.minneapolisunions.org/mlr2019-07-26_shakopee_strike.php [accessed on April 12, 2020].

- At the *participants level* (those who help long-term decision-making), new products and processes are designed for a human task force that is augmented by, or in some cases displaced by, robotic systems.

As we will explore in later chapters, the designers, managers, and researchers at the *participants level* will be essential for maintaining human, ethical values in the workplace. They will redefine the tasks and skills needed by the human labor force (at the implementation level) and will design the objective functions and data that are used to train robots and other AI-driven processes.

Summary and Conclusion

In this chapter we have explored how the workforce was restructured in the Information Age in order to support a networked, knowledge economy, and we have examined the different ways of working and interacting with robots.

Unlike previous technological revolutions, information technology tends to devalue jobs that are physically and cognitively repetitive but cannot yet be automated. The flip side of this tendency is that information technology increases the value of jobs that focus on social networking, process design, and creativity. However, even these lucrative, creative jobs are at risk. Robotics and AI will transform research, design, and project integration jobs, and they will increasingly participate in strategic decisions at the highest levels.[74]

[74]This is already happening—Deep Knowledge Ventures (DVK) appointed an AI algorithm, vital to its board with the right to vote on important decisions. Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. California Management Review, 61(4), 66-83. Also see Burridge, N. (May 10, 2017). "Artificial Intelligence Gets a Seat in the Boardroom: Hong Kong Venture Capitalist Sees AI Running Asian Companies within 5 Years," *Nikkei Asian Review*, https://asia.nikkei.com/Business/Artificial-intelligence-gets-a-seat-in-the-boardroom.

These conclusions are reflected in the Future of Jobs Report 2018 developed by the World Economic Forum (WEF). Cognitive tasks, from routine data processing to complex decision-making and coordination, are shifting from human to machine labor. This does not mean that jobs will decline (at least not initially) but rather that robots and intelligent automation will be involved in more and more of the tasks associated with those jobs.

There is no single way robots and AI will transform work. In some cases, human performance will be augmented through wearable computing and remote-controlled robots. Remote-controlled surgical robotics, for example, has is benefits and drawback, as discussed in later chapters, but it is now part of the healthcare system and it will continue to evolve. The research challenges for this style of work transformation focus on optimizing the user experience: the human operator needs to feel situated and in control.

In other cases, AI and robots will increasingly take over decision-making and perhaps executive management functions. This might create a utopia in which humans enjoy more discretionary free time, or it might create dystopia in which humans are subjugated. When workers are replaced by autonomous machines, instilling ethical reasoning and human values into their design and operation becomes the principal research challenge. We have already witnessed the dangers of allowing AI programs and training data to make consequential decisions that reflect and thereby repeat human biases and prejudices. "Too many executives have chosen to displace workers rather than think through how technology and humans can work together symbiotically."[75]

[75]L. P. Willcocks & M. C. Lacity, (2015). Nine likely scenarios arising from the growing use of robots. https://blogs.lse.ac.uk/businessreview/2015/09/29/nine-likely-scenarios-arising-from-the-growing-use-of-robots/ [accessed on April 9, 2020].

Thinking through how humans and robots can work together as partners is the third way in which AI and robotics will transform work. Humans and collaborative robots (cobots) will partner to form a symbiotic relationship, like the sort of relationship humans have formed with work animals, especially dogs, albeit in this case, robots may eventually become equal partners. In this last form of work transformation, research into team coordination, collaboration, and relation-making becomes critical.[76]

Just as ergonomics was developed to make tools and materials (e.g., containers) easier for humans to use (physically and cognitively), the next generation of production tools and materials will need to consider the limits and abilities of both humans and robots (although the latter may be codesigned with the rest of the production environment).

In conclusion, over the next decade, the human workforce will shift away from implementation (except for expert craftsman marketing "made by human hands" products) and toward participation in decision-making and robot team supervision. Some technologists such as Martin Ford and business leaders such as Elon Musk believe that if left unchecked, robots will eventually dominate all aspects of human labor, including executive decisions and creative research and design of work. Others argue for a more symbiotic relationship in which human and collaborative robots (cobots) are partners.

We take the view that humans and cobots working as a team are typically superior to machine-only systems, and that human-cobot systems are a desirable social goal. This is a technology challenge and design

[76]In all cases, ethics and bias are research challenges for the design and operation of these systems, and this has implications for work during the next decade. To ensure that robots do not discriminate against certain groups or show preferential treatment to some groups, the workforce of executive management, researchers, designers, integrators, and implementers must be diverse, reflecting the diversity of society.

goal, reminiscent of Schumacher's *Small is Beautiful*[77]: Efficient, low-cost systems can be designed so that tasks can require human craft as well as robotic capabilities. How this might be achieved—the research challenges and current state of the art in addressing these challenges—will be addressed in the remainder of this book.

[77]Schumacher, E. F. (1973). *Small is beautiful: a study of economics as if people mattered*. Vintage. Schumacher's essays provided a much-needed critique of western investments in the developing economies and the "bigger is better" approach. He advocated the use of small-scale technologies that were appropriate to the situation, decentralized, environmentally sound, and consistent with human dignity and empowerment.

CHAPTER 2

Technology Definitions

Level Setting for the Rest of the Book

Before discussing technology challenges, it is worth looking at how artificial intelligence, machine learning, reinforcement learning, and neural networks relate to each other, what are they, and what is their relationship to automation and collaborative robotics.

These definitions are for related technologies with some foundational technology; as shown in Figure 2-1, cloud computing is an underlying enabler, but may not be used in all circumstances. It is also a well-established technology that needs little or no explanation. It is included here because of its capability to flexibly manage the massive amounts of data that form the basis for accurate AI and automation.

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_2

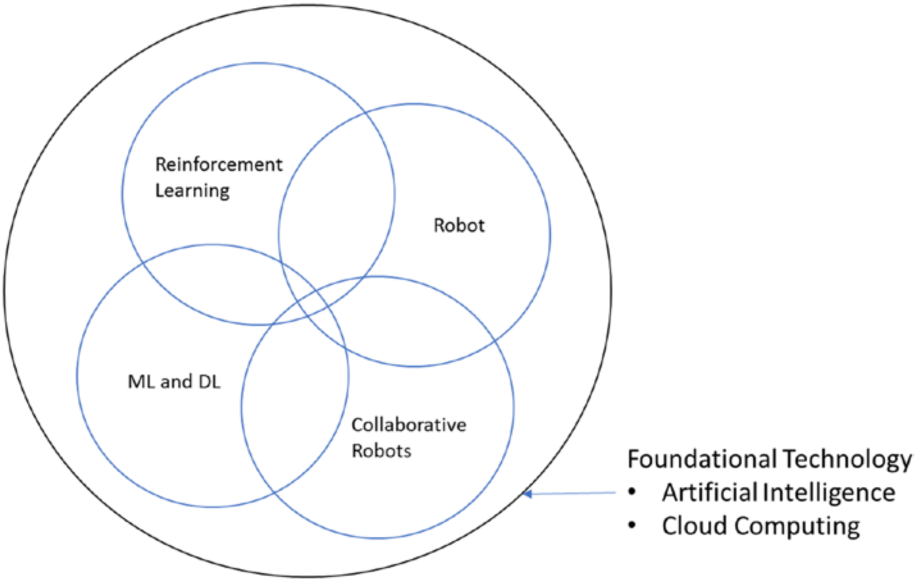

Figure 2-1. *Initial definitions and interrelationships*

The major enabling technology for automation and robotics is artificial intelligence (AI). This is such a large domain that it includes several subdomains that are significant in the context of this book.

Definitions

The following are not exhaustive definitions, but give pragmatic descriptions and context of the technology.

Artificial Intelligence (AI)

Britannica Online[1] defines AI as the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. This is a sufficiently broad description that covers a

[1]Encyclopedia Britannica online: www.britannica.com/technology/artificial-intelligence

huge domain of technologies and techniques. New subdomains have been added over the last 20 years, for example, deep learning that emerged to support big data.

Earlier AI tools were only partially successful working on small data sets. They were also subject to inflated expectations of the outcome of an AI application. In several cases the highly priced tools were sold as a “silver bullet” solution to the problem of too much complex unstructured data for the existing analytical tools. Some implementations were successful but not adaptable and scalable enough.

One of the issues behind AI’s problems, in the 1990s, was the lack of good data analysts and AI expertise outside of the academic world. The lack of good data is an equally important factor in the failure of earlier AI solutions. Good data not only refers to the volume of data but also the quality. Data that has a bias will produce biased models and conclusions. Some data played through a predictive model can’t give a better than 50% level of accuracy. In the late 1990s there were several attempts to produce predictive models for use in retail stores and insurance companies. The retail store project was focused on handheld scanning devices. The concept was simple. After the customer had scanned several items, the scan was used to consult a neural network to predict the customer buying pattern. Successful prediction could then prompt a message to the scanner to try and sell other goods based on their purchases. For example, if a customer had purchased meat and charcoal, a message could be sent telling them that barbecue sauce is in aisle 20. Despite gathering a suitable amount of data whenever the model was checked, it could only give between 47% and 53% accuracy. This was not good enough to justify an investment by the store owners. The salespeople who contacted the retail store owners had given a really unrealistic view of the capabilities of proposed AI solution that could not be met. Hype and wild statements about the potential values of a solution led to unrealistic expectations and the project was cancelled. In the same way that service-oriented architectures were a technology looking for a solution, so was the AI of the 1990s. Automation and robotics would not be as pervasive as it promises to become without the current advances in processing power, cloud computing, and deep or machine learning that go beyond anything that was proposed in those times.

Neural Networks

The Collins online dictionary defines a neural network as a program or system that is modeled on nature, more specifically the human brain, and is designed to imitate the brain's method of functioning, particularly the process of learning.[2] A neural network, often called an artificial neural network (ANN), consists of vast numbers of simple nodes as seen in Figure 2-2.

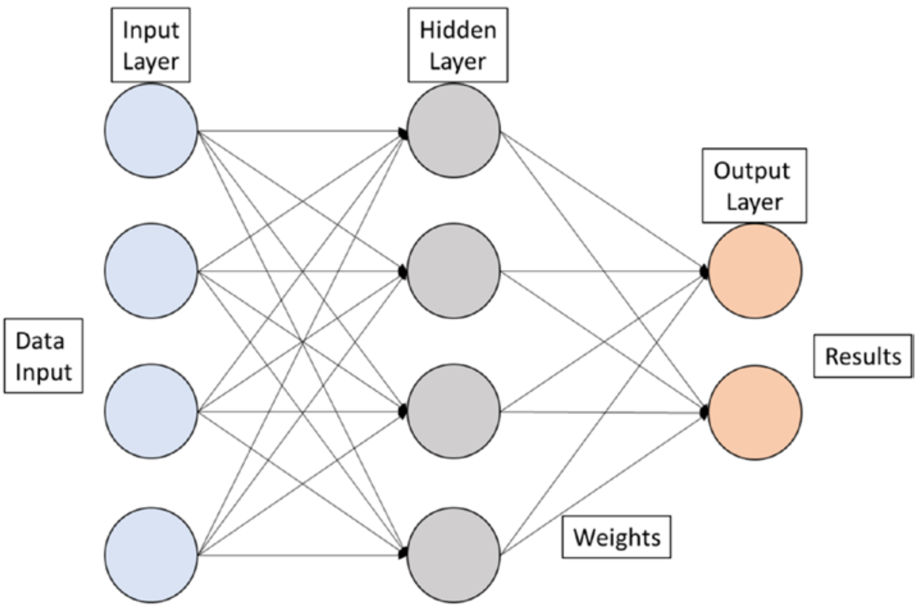

Figure 2-2. *Neural network topology*

[2]Collins Online Dictionary: www.collinsdictionary.com/dictionary/english/neural-network

Each of the nodes in a neural net is densely connected to other nodes in the network, somewhat mimicking the way that a human brain is wired. In an artificial neural network, there are several layers of nodes. In Figure 2-2, the Input node, on the left, is for raw data, while the output layer is responsible for computations and presenting the results to the outside world. The hidden layer in the figure has no contact with the outside world, hence hidden. The hidden layer transfers information from the input nodes to the output nodes and carries out computations. There are other topologies where there are more hidden layers but there is no need to go into that level of detail. All of these nodes are densely interconnected, and the connections are given a numerical weight.

When the neural network is fed with known labeled data, it is called supervised learning, because the network is trained by labeled input and output data prior to being fed raw input data. Weights are the strength of connection between the nodes in a network. When new raw data is fed into a trained neural net, it first establishes that the data is within the bounds of the training data, and if so, it then calculates the possible output result.

Neural networks have gone in and out of fashion since they were discussed as early as 1944. Their emergence now is based on new algorithms, techniques, available data, and the huge increase in processing power offered by graphics chips, among other technologies.

Machine Learning/Deep Learning

Neural networks, a technique often used today in machine learning (ML) may be based on supervised or unsupervised learning. Unsupervised learning comes into its own when nothing is known about the data; there are no labeled input and output data classifications. Machine learning (ML) is a method of data analysis that automates analytical model building. It is a branch of AI based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention. During the learning process, the analytical model generally requires no additional programming by a human (the weights between connected are automatically adjusted). However there is considerable programming work in gathering data and passing it to the ML algorithm.

Deep learning is a subset of techniques that use neural networks and is often used on very large amounts of data, sometimes called big data. Big data is used as a term for massive volumes of both structured and unstructured data that is so large it is difficult to process using traditional database and software techniques. Big data is data that changes too rapidly, or that exceeds processing capacity available or just too big to be managed.

Deep learning filters data by using hierarchical layers of neural networks to process the data input. The output for one layer becomes the input for another layer until the result is available. For example, in the computer vision domain identifying a face in an image, the deep-learning solution can use that image as the input represented by a matrix of pixels. The first layer would then encode the edges and compose the pixels. The next layer might compose an arrangement of edges. The next layer might encode a nose and eyes. The next layer might recognize that the image contains a face, and so on. This approach, many layers deep, gives the name deep learning.

This type of machine learning can process such vast amounts of data that a human data analyst may take decades to process and understand. A good example of the power and data demands of deep learning is the speech recognition functions of smartphones and digital assistants.

Reinforcement Learning

Reinforcement learning is a form of machine learning that can develop rules to solve problems on its own. Reinforcement learning is an autonomous self-teaching system that essentially learns by trial and error. The goal of this learning method is to maximize the best possible outcomes. A robot can grasp a glass and if it grasps too firmly the glass will shatter. The robot will then try grasping the glass with less pressure, and if it drops the glass, it can then try to increase the pressure slightly until it can hold the glass without crushing it or dropping it. The robot has used reinforcement learning to arrive at a good outcome and produce a desired approach. The robot gets feedback from the external world, and the type of feedback, positive or negative, will condition the next actions of the robot until it arrives at and remembers the best solution. This is one of the fundamental building blocks of robotics.

Robot

This is a difficult definition because there are so many alternatives and even roboticists cannot agree completely. The Institute of Electrical and Electronics Engineers (IEEE) has a basic page for learning about robots and it has a generic definition that will suit our purposes. The definition says that a robot "is an autonomous machine capable of sensing its environment, carrying out computations to make decisions, and performing actions in the real world."[3] This definition is broad enough to cover industrial robots, household robots such as the Roomba, and collaborative robots that work and interact with humans to complete a set of tasks. Because this book is looking at collaborative robots, this definition will suffice.

Collaborative Robot (Cobot)

A collaborative robot or cobot is a robot that can safely and effectively interact and cooperate with humans in performing a variety of tasks. Collaboration can take many forms, from the simple, such as passing material for construction, to a human or working in a team of humans and robots to accomplish a more complex task. Those tasks can use the robot's strength or reach to compliment a human's capabilities. An important factor in any situation involving a cobot is colocation with a human to complete a task. Unlike industrial robots, in a cobot's collaborative world,

[3] IEEE Learn: Robots https://robots.ieee.org/learn/

the cobot is not fenced off or in a separate space, so the humans and robots have to interact safely. A simplistic rule that tells a cobot to stop when it detects a human will not work since the robot may be carrying something to the human for the next stage of a task.

Automation

Automation is defined in the Cambridge English dictionary as the use of machines and computers that can operate without needing human control. Automation can carry out physical tasks such as picking goods for dispatch or nonphysical tasks such as ranking loan applications by using a software robot. We will refer to physical work being done by a robot as robotics and intelligence led work such as calculating and approving a mortgage being done by a software robot as automation.

Some of the earliest information technology automation was the movement of clerical tasks such as ledger entry and reporting into IT applications. At this early time, there was still a need for clerks to enter data manually, often using a key punch bureau, but the reconciliation of different ledgers and consolidated reporting was automated. This led to clerical ledger maintenance jobs being lost but led to the rise in new data entry jobs. Ledger maintenance would rely on a set of rules that could be coded into the maintenance application. Frequently the rules were difficult to change requiring programmers to spend time updating an application.

Techniques such as parameterizing many of the rules would enable easier changes. Over time these techniques were improved, and even newer ideas were developed with advances in programming languages and databases. These improvements were also matched by an increase in computing power. Microcomputers and personal computers moved the data processing from the large centralized mainframe to smaller machines that were capable of taking the load and moving it into departments or even offices. More tasks became automated and application configuration provided flexibility and automation took another leap forward. At this time, one of the most significant automation tools revolutionized the work of accountants and financial officers and did it without computer engineers or programmers. The spreadsheet that was developed for personal computers, VisiCalc,[4] enabled fast and accurate updates to financial models and automated a process that was often carried out by hand with pencil and eraser. Its appeal was based on the familiarity with the spreadsheet modeling and in the immediacy of entry and calculation. The successors of VisiCalc have led to the situation today where most people with a home computer or a phone can carry out sophisticated operations without much instruction. The personal computer is now ubiquitous.

Automation of physical tasks accelerated with the first industrial robots implemented at General Motors in 1961.[5] The use of the production line in the automotive industry with well-defined repetitive tasks was an important factor in the automation of vehicle manufacture—massive automation in the automotive industry combined with a realization that well-defined repetitive tasks can be automated and can run 24 hours a day 7 days a week.

Programming for the early industrial robots was complex and used mathematics to calculate the angles of various joints stored in a teaching phase and replayed in operation. While the accuracy of 1/10,000th of an inch was essential and possible, changing the programming to handle a different design was time consuming, often taking days, and complex requiring a new training process.

Automation either in the physical or the software domains have exposed a number of challenges that are increasing in significance as the scope of automation and robotics grows. Challenges such as synchronization of automated actions and error management, including the prediction of possible errors, combined with risk management and

[4] History of Computers and Computing, Birth of the modern computer, Software history, VisiCalc of Dan Bricklin and Bob Frankston (n.d.). Retrieved April 9, 2020, from https://history-computer.com/ModernComputer/Software/Visicalc.html

[5] First Industrial Robots: https://ifr.org/robot-history

enhanced decision-making are all trying to simplify the interaction between IT, robotics, workers, and customers. Some of these challenges need to be met and overcome before robots become autonomous. Simple decision-making has already been mentioned where an industrial robot makes a simple decision to stop when a human enters the security zone. Advanced decision-making can allow a robot to continue moving once it has detected the human by constantly reviewing their relative positions, speed, and direction of motion. Without these challenges being met, automation and collaborative robotics will not be able to fulfill their promise.

Rules That Don't Work, Bots, and Chatbots

Software automation is evolving and the impact that it will have on the workers of the future is already being felt. Early software automation through the use of business rules is already in place in many organizations. Poorly written business rules can cause disruption and errors in any organization. We are all familiar with computer errors that can register a default on a payment when there has been no default or that credits to a utility account are not made in a timely manner. Following these problems, the customer frequently has to spend a long time on a telephone to resolve the situation. This type of error is often generated by a disconnect between the line-of-business (LOB) team who understand the business model and its rules and the IT department who has to implement the rules without understanding the business model. It is possible to say in this circumstance that rules don't work.

A continual challenge that has not been completely conquered is the differences in understanding between line of business and the IT department. Often, still, the business complains that the IT department does not understand them, and the IT department complains that the business does not define their processes fully enough. There have been many attempts to solve this disconnect but it is still present.

A workshop in the late 2000s between lawyers, business people, and IT experts was held to try to outline and codify the borders between regulations from governments, lawyers who interpret the regulations, businesses that must comply with the regulations, and finally the IT world that has to write software to manage the compliance.

At this workshop a lawyer stood up and said that the law was very simple and easy to follow, but business and IT kept getting muddled. The business people said that the law was too complex for them to be able to guarantee compliance and lawyers need to speak plainly. The IT people then stood up and said that they can code for anything, but the lawyers and the businesses were not being clear. There is little evidence that this disconnect has been substantially solved but there have been a number of attempts.

In the 1990s, a Structured Query Language (SQL) was developed as a technology solution to allow business users to query data for reports and results without asking the IT department, needless to say the IT department had to develop the queries and supporting applications but there was still a disconnect. A subsequent technological solution, the business rules engine, was promoted as a more effective method of encoding business rules solutions without resorting to writing code. These rules engines are still being used, but the rules are mostly written by software engineers not business people and again we have the disconnect.

A significant advance in recent years is the growth of robotic process automation (RPA). This is software robotics applied to existing automatable processes and it is discussed in more detail in a later chapter. Instead of writing rules, the RPA tools learn how to follow a process by copying the actions of a user of all the applications in a business process. The user is a business process expert not an IT expert, so at last the disconnect between business and IT is minimized. RPA is currently popular because it can yield cost savings and productivity gains; however, there should be some caution. Several analysts are suggesting that RPA should not be implemented by IT departments because of the history of delivery failures of IT projects.

While RPA allows the automation of a business process by learning actions, it is not able to take complex decisions. Chatbots or conversational systems are capable of following a process and using AI-based decision-making to respond more intuitively to humans. Software automation has evolved to such an extent that some of the chatbots used in call centers produce interactions that are hard to differentiate from human-to-human interaction.

Robots, Collaboration, and Collaborative Robots

Robots and robotics are commonplace thanks to the entertainment media as well as the scientific and business communities. Few can imagine a future where robots are not a major factor in all walks of life, and although the media tends to anthropomorphize representations, pragmatic robots tend to have shapes that match the function and purpose they have been designed for.

There are some interesting thoughts regarding the repurposing of specialist robots to new tasks. Industrial robots have been mentioned before and play no part here, but personal and business environments form part of a collaborative environment for robotics.

One of the best known and earliest home help robots is the Roomba. This autonomous room cleaning robot was introduced by the iRobot Corporation in 2002 and has gone through a number of iterations, making it easy to recognize.[6] In the future domestic robots will continue to aid and assist people to achieve a good quality of life.

[6]iRobot's Roomba: https://web.archive.org/web/20120103091646/http://www.irobot.com/sp.cfm?pageid=203

One of the more compelling uses of automation and robots is in the healthcare industry. Many organizations are developing solutions that can provide care in the home. One of the case studies later in the book, the ENACT project, features a partner delivering tools and automation for healthcare in the home.

Healthcare and health monitoring delivered successfully can enable the infirm or elderly to live in their own homes for longer than the current technology and level of care permits. This has many advantages both in well-being, finance, and use of medical resources. The architecture of this case study is discussed more fully later in the chapter and illustrated by Figure 2-3. Some of the key points are as follows:

- The resident is emotionally attached to their home and is more relaxed in there. If you ask elderly people their view, it is almost unanimous that they want to stay in their own home.
- The costs of maintaining someone in their own home are high, but still cheaper than staying in a hospital. And again, ask a patient what they want most from their treatment and it is "to go home." This applies to most patients, but is emphasized by elderly patients, in our experience.
- Hospitals are not good places for long-term stays.
- The use of nurses for nonmedical activities is not always an effective use of resources.

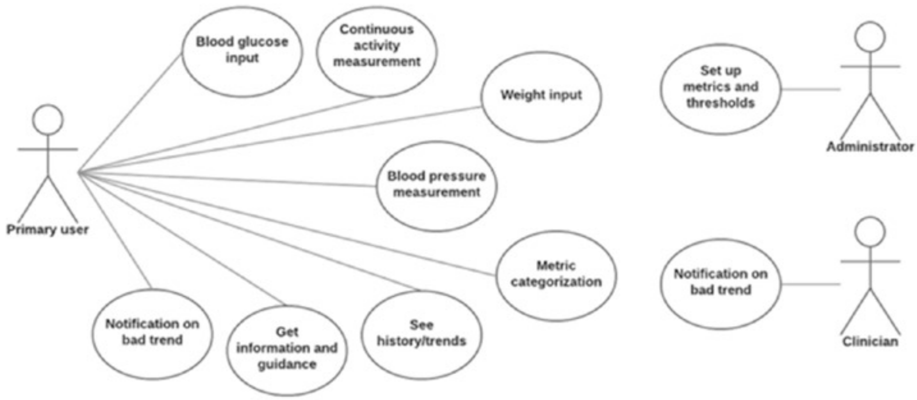

Figure 2-3. *ENACT healthcare solution*

Monitoring of the resident, their environment, complementing the resident's activities, and socializing online or in their own home by using technology will, in the future, remove some of the societal pressures of managing an aging population. Not all of this work should be outsourced to automated tools and robots—there will need to be more consideration paid to empathy and socializing with friends, relatives, and caregivers—however there is scope in this case to supplement care of the aged with tools that can think, to improve quality of life. Many of the sensors that are reporting on the elderly person's environment are not typical networked computers but low-powered devices connected either to the Internet or to a small local computer. These devices are said to be at the edge of the network. Automating risk and decision-making at the edge of the network is an area that will be discussed in more detail later in the book.

There is a difference between using robots to complete a simple task and the more specialized automated healthcare. Automated healthcare will use an integrated solution including robotics, sensors, activators, and monitoring tools.

Another area of potential growth in automation and robotics is the supply chain. Organizations are increasingly using robots as part of the supply chain and many cases of robots in warehouses are quoted in the press. Amazon, for example, uses robots in large action spaces. These robots are moving in a known space in the warehouse. That space is fixed and doesn't change without corresponding changes to the robots, humans, and any markings used to delineate the space. This reduces the need for complex programming and remapping of the space. These are relatively simple robots that perform well-defined tasks within bounds (the warehouse) and with people aware that they are in the robot's space rather than a shared space.

Collaboration is one of the more difficult aspects of human/robot interaction but has the potential to change the relationship between robots and people by introducing shared spaces and shared tasks. This will require a high degree of communication, a jointly understandable view of the world, and some sophisticated policies and metrics. There is the potential that collaborative robotics may even dilute or negate any Luddite tendency by showing humans and robots working in partnership rather than as adversaries.

Another important feature of collaborative robots is the level of instruction and the interpretation of those instructions. An industrial robot will require complex programming and instructions, often achieved by moving the robot appendages to a particular location and then programming the action of the appendage. A collaborative robot will be expected to cope with simple instructions in the vein of "Move that chair out of the way" or "Clean up that mess." Interpreting these instructions and doing so in safety and an environment that understands risk management and mitigation will place extra demands on the processing, communication, and risk management capacity of each entity in the domain.

Smart Buildings As Robots Without Arms

Smart building technology has been in use for many years, but technology that creates individual environments can be said to be in its infancy, handling light, heating, and other environmental sensors. There are some leading-edge buildings that are moving forward by integrating internal booking systems and motion sensors to be able to tell if a meeting room is occupied, to remotely diagnose equipment and environmental failures, and to check the diaries of room booking and employees to resolve any issues. In this way buildings are becoming smarter and will eventually incorporate collaboration between people, machines, buildings, and software, even extending to the environment outside the building.

Work in the future will include interacting with many devices that will streamline and remove errors and risks in a business environment. An example of an “Empathic Building” from Tieto corporation will indicate progress toward the truly smart and interactive building. As part of this, another impact on work and employment in the future will come from developments in root cause analysis of error or failure conditions. Some of this will come from the root cause analysis of issues in a large, edge of the network environment which will be discussed in more detail later.

Smart buildings are not the only special case. Autonomous vehicles have many of the characteristics of a smart building or a collaborative robot, but they are a mode of transport. Like a smart building they have software automation and sensor/activator integration. An autonomous vehicle collaborates with a human (the passenger) and is designed to operate in a dynamically changing environment, making decisions based on the vehicle’s view of the environment, the safety policy, and the instructions on the journey. Autonomous vehicles also have to interact with the external environment, with people in the roadway, temporary road signs, and the weather.

They exhibit many of the complexities of a collaborative robot, but in particular the mapping and physical location of the vehicle are demanding because of the number of other vehicles and the ever-changing road environments. This is a domain that has a clear impact on the future of work and is likely to be one of the most recognizable factors affecting employment. Robots in warehouses are common; however, introducing autonomous vehicles into the supply chain at both delivery and fulfillment will create seismic changes on employment in those areas.

Research Progress

Scientific research is a precursor to technological advances. Examining scientific research in any field is a good indicator of both progress and direction in that field. The scope of robotics, and automation research is huge and covers many areas from construction, programming, developing new models of communication and movement through to policy and decision making. The authors have been engaged in a number of research projects and we have selected a few areas that we believe are important foundational technologies for collaborative automation and robotics. We introduce some of these domains below and they are expanded in the corresponding chapters in the rest of the book. Much of the research we discuss is promising and has been demonstrated in laboratory conditions. Other domains we discuss are maturing but missing some crucial element.

Data Fusion

To collaborate together, humans and different styles of robot need a common view of their working environment. This will require a number of tools and techniques such as data fusion and computer vision. Data fusion is explored in this book as it leads to a common model of the environment available for both robots and humans alike.

Many researchers agree that the most serviceable definition of data fusion came from the Joint Directors of Laboratories Workshop.[7] This defined data fusion as "a multi-level approach dealing with the association of data and information from single and multiple sources to achieve refined position, identify estimates and complete and timely assessments of situations, threats and their significance." In a world that is increasingly being defined by sensors, the number of errors and conflicting data/information is expected to grow. Robots working in a shared space need to have access to near real-time assessments of a situation that can ensure safe and accurate performance of tasks.

In an article for the *Scientific World Journal*, Federico Castanedo reviews data fusion techniques.[8] This review has led us to the conclusion that the most effective data fusion architecture is likely to be a decentralized one. Different architectures are defined later in the book.

As computing power increases at the edge of the network, it is likely that robots with autonomy will increasingly take decisions and analyze risks using a decentralized model although there will be some level of centralized control and communication. Even limited situational awareness in a robot can only be achieved by having a common model with the other actors in the shared space. This is included as a challenge since some of this work is in its infancy. To establish progress and future plans for this research, we will conduct interviews with two researchers. Professor Moncef Gabbouj of Tampere University department of signal processing has been working on data fusion for some time and his interview is recorded and discussed later in the book.

[7]JDL, Data Fusion Lexicon, Technical Panel for C3, F.E. White, San Diego, California, USA 1991.

[8]Federico Castanedo, A Review of Data Fusion Techniques, The Scientific World Journal, Volume 2013, 2013.

Real-World Common Models

The goal of data fusion is to generate real-world common models that can be understood or interpreted by both humans and machines collaborating in a shared space. A common model in this context is a model using identical data that will be kept current and able to be interpreted by any of the actors, human or robot, engaged in a collaboration. Data fusion will take data from large numbers of different sensors, activators, and cameras and merge them into a common model. Computer vision research and visualization of large, complex data are important inputs to this common model and are treated separately here, since they are significant research topics in their own right. Computer vision is advancing but is not enough on its own to develop a good enough model for collaborative robotics.

Project ENACT

The healthcare case study in ENACT is being developed by Norwegian startup company Tellu IoT AS.[9] Figure 2-3 showed the range of data and communications used in the case study to deliver practical proof points for the research. Tellu's objective as part of ENACT is to combine simple sensors and actuators to manage a home for an elderly or less able resident. If a house plant needs watering, a sensor will detect this and send a message to the resident or their caretaker, so forgetting to water a plant should be a thing of the past. If this is integrated with more complex medical sensors, motion sensors, and other sensors, this can ensure that the resident can stay in their own home for far longer than is currently possible.

At first sight the ENACT healthcare case study has little to do with a collaborative robot; however, problems solved in the ENACT project will help solve similar problems in a cobot.

---

[9]Tellu IoT AS, www.tellucloud.com/

The Internet of Things (IoT) is a system of computing devices, mechanical and digital machines, sensors, and actuators that can be uniquely identified. They have the stability to transfer data over a network without requiring interactions with humans or other computers. A cobot consists of a body or frame, a control system, manipulators, and some means of travel.

The architecture of a cobot, with networked sensors and actuators communicating with a control system and also through Internet connections communicating with the wider world, has similarities with the house in one of the ENACT case studies. In this house sensors and actuators placed around the house are networked with a control system, the walls of the house are the body, and there is no need for the house to move so there is no means of travel. Trustworthiness is an important factor in both architectures.

The ENACT project will research into the development of smart information systems, similar to those needed in a cobot. Smart information systems will take into account security, privacy, resilience, and robustness in both the IoT-based solutions for the healthcare system and the internal networking of a cobot.

Collaboration and Policy

Discussions on collaborative robots and their impact on the future of work often hinge around the safety aspects of having mobile machines inhabit the same space as a human. There is a general fear of "robots out of control" and safety policies have to be one of the most important and visible policies. Robots that can only be used in a safe known space, such as static industrial robots, are unlikely to have a more major effect on the future of work than they already have. All evidence indicates that mobile, autonomous robots, cobots, and software automation will become more pervasive. Safety must be at the front of any discussions about collaborative robotics and the most important factor that needs to be resolved as soon as possible. The predicted effect of collaborative robots on the future of work will only be possible if human collaborators and workers feel safe.

Collaboration will require communications to be integrated and understandable by all the actors in a scenario including humans, cobots, other robots, and automated processes. The policies and processes must be flexible and easy to implement and need to solve increasingly difficult problems.

One of the most difficult domains in robotics currently receiving a lot of attention (and finance) is grasping or picking up objects. Humans from a very early age learn to grasp items although that is mostly done by trial and error. A good example is a child grasping an egg. It is difficult for a young child to know how much pressure to apply to an egg to successfully lift it, but not crush the shell. Robots have a similar problem. However, there is a related problem in collaborative robotics: the handover. If a robot holds an object and passes it to another robot or a human, how does the first robot know if the second actor has a good enough grip on the object?

Another problem is that as some stage in the process two actors hold the object at the same time and there has to be a communication that indicates that one of the actors is ready to let go and the other actor is ready to hold the object on their own. This and the grasping/picking up problem will have to be solved for future progress in collaborative robotics. Perhaps part of the solution could be a redesign of the object being grasped. For example, a beer company has redesigned their glasses without handles so that when more than two glasses are grasped by two hands the glasses lock together enabling that person's two hands to grasp three, four, or even five glasses without dropping them.

Once robots come out of the factory floor into the store or open spaces, the level of risk increases. Mitigating risks is a valued strategy to reduce or manage risks; however, the risks must be identified. The use of continuous risk assessment and management of risks in collaborative robotics is a relatively new field. In any situation humans and cobots will have to continuously assess the risk in that situation and resolve or mitigate that risk by management actions. There is a body of research into risk management and decision-making in the ENACT project that will also be significant in the future.

Initial thoughts on the future ethical conundrums of robot and cobot rights throw out some interesting questions; for example, when is a cobot allowed to move despite a threat to the safety of an actor or how can a cobot discriminate between a benevolent actor and a malicious actor? There are also ethical questions and concerns regarding the military use of automation and cobotics. All of these and many more issues will have to be addressed before automation and collaborative robotics takes its place as an integral part of the work force of the future.

Summary and Conclusion

There are a number of significant topics that need clarification before the technology and policy discussions in the rest of the book can be digested. We have noted simple descriptions of key concepts such as artificial neural networks, deep learning, and data fusion. The descriptions are introduced as level-setting concepts and are tailored for use with this book. We have followed technical definitions with some of the basic concepts including robotic process automation, conversational robots, and collaborative robots.

Research progress is one of the strengths of this book since it will give a picture of the progressing development of solutions. These solutions to barriers of adoption for collaborative robots and automation are the gating factor for fully autonomous collaborating robots and humans. Finally, we introduce a number of interviews with lead investigators on data fusion and healthcare at the edge of the network to give the latest views on technology research in their area. This is a fast-paced world and advances are being made and announced frequently. We are also met with a weekly, sometimes daily, article on the positive, negative, or occasionally neutral effect that automation and collaborative robotics will have on the workforce.

There is no doubt that work in the future will use automation and collaborative robotics, but the timescales are often unrealistic. We hope to show that research is being conducted to solve some of the really large problems, and we have presented a number of research projects that directly or indirectly contribute to solving some of these problems. This chapter completes the book section “Preparing for the Future of Work” and is a level set for the next book section dedicated to what robots are doing called “Robots Are Working.”

PART II

Robots Are Working

CHAPTER 3

Robotic Process Automation

Is This the Real Job Killer?

Although the subtitle of this book is "A Guide to the Future of Work," this chapter will explore the challenges of automation and some solutions that have a legitimate claim to affect the present and the future of work. As we will show in this chapter, automation is not a series of distinct moments but a continuum of strategies and technologies that intend to increase levels of automation and improve the efficiency and effectiveness of business operations. Some of the strategies and technologies we discuss are already being deployed; others are only at the initial stages but will have an impact on who does what work in the future. Increasing automation has been a goal of information technology since the earliest attempts to use computing to handle vast amounts of data and repetitive tasks. If we look at some of the significant places on the continuum of automation, we can gain perspective.

In the 1940s, Colossus, one of the earliest computers, automated the reading and comparison of vast amounts of encoded cipher text that helped humans to decrypt secret messages.

At the next significant point, in the 1960s, along the continuum record keeping and particularly business ledgers led to a need to further

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_3

automate customer service interactions and decision-making. Much of the automation at this time was concerned with making the automated tasks faster to increase a business's capacity. Office computers became smaller and automation delivered more finely defined and efficient business responses. On one occasion one of the authors worked with a business division that was part of a large conglomerate and had to report their financial status in a head office briefing every Friday. The data preparation for the company computer bureau required the efforts of two people for four days, finally sending all the data to the bureau for a report to be generated for the Friday briefing. Installing a small computer, an Apple IIe, meant that not only can the data be entered into a ledger as it was gathered, interim reports could be completed and the report for the Friday briefing prepared, reviewed, and sent without using a bureau at all. This allowed the staff to be redeployed on more business-focused worth and meant that the financial status of the division was known on a daily basis. Automating the task locally generated better business practices and efficiency and was soon copied across the whole organization. In most businesses staff are now at a stage where they can get almost all the information that they need instantly. Automation can result in a reduction of human effort in repetitive tasks, allowing them to be redeployed into more creative work. It also reduces human error in repeatable tasks. Humans have difficulty working on repetitive tasks for long periods of time, becoming bored and more prone to mistakes.

At the current time many organizations are evaluating or implementing a technology called robotic process automation (RPA). RPA takes automation one step further by taking previously difficult to automate tasks and automating them. RPA uses software that can be trained to automate a business process rather than a task, removing humans from a larger portion of the business process rather than just taking a task from a human but still requiring the human to operate the business process.

RPA has a lot of potential but there are also some challenges. We will review the challenges of implementing RPA as well as the potential for RPA to exploit new opportunities for that automation. RPA has the enticing prospect of automating more complex repetitive tasks that are currently fulfilled by humans. RPA does this by bypassing the need for programming to effect new automation, learning by watching the process being followed through the graphical user interface. An RPA tool, which is a metaphorical software robot or bot, will watch a human perform a series of actions on their computer that follows a business process. These actions are recorded and the bot can then execute these actions. Actions can be complex or simple, for example, open up a customer account search, conduct the search, and copy the customer number into an invoicing page. There is no programming because the bot learns the process and can relearn a changed process and repeat it without a human operator. Most importantly it does this at an industrial scale. RPA offers the potential to streamline practices and cut costs by running applications 24x7 with far fewer humans involved in a process. When it comes to overall management of a business process, RPA does not stand on its own; organizations will need other tools such as business process analysis (BPA) or business process optimization (BPO). RPA cannot itself change process or optimize it, RPA only executes a process. BPA and BPO will enhance the performance of an implementation of RPA by ensuring that the process being automated is optimized and delivering the required results.

RPA is already being implemented in many businesses. The value of increasing automation is generating an interest in using AI for the next-generation RPA. Machine learning will add intelligence to further automate process flow and play a greater part in decision-making. There are other factors that businesses invest in to automate processes including offshoring and outsourcing tasks. The impact of RPA on these other factors will be discussed later in the chapter.

RPA, offshoring, and outsourcing are different approaches to addressing automation. RPA is using technology to execute processes without human operators. Offshoring and outsourcing are approaches that move the responsibility for process automation out of the business itself to a third party. The third party can decide if they use humans, for example, in an offshore call center, or use automation tools like RPA. There are cost implications for both the process owning business and the third party. In third parties the debate may rest on the cost comparisons between human and software robot operations. In one such discussion with a third-party supplier, they suggested that software robots have many advantages, but the software is expensive and the third-party supplier worked in an offshore environment that had low-cost labor. They commented that it was cheaper to hire 100 staff to automate a process than to buy the software that would replicate the work of 100 staff.

Increasing automation of business processes for cost or efficiency reasons is a primary operational goal for business. The benefit that is often overlooked in a cost-benefit analysis of automation is the value of increased and manageable regulatory compliance. A good example of the importance of automation is the scramble to become European Union General Data Protection Regulation (GDPR) compliant.

GDPR introduced a new concept. It applied to the European Union (EU) states, of course, but it also had an impact on businesses in states outside the EU. These businesses would have to comply, wherever they were based. Failure to comply could result in large fines of up to 4% of a company's global turnover, for example, Google was fined $50 million for a breach of GDPR.[1] Penalties can also include sending the accountable investment became even more important when the method of calculating fines for noncompliance was evaluated and indicated that it gives considerable powers to regulators to punish companies. Individual states

[1]Fox, C. (2019, January 21). Google hit with £44m GDPR fine over ads. Retrieved April 2, 2020, from www.bbc.co.uk/news/technology-46944696

in the European Union also have the power to make noncompliance a criminal offence, opening the possibility for custodial sentences for the accountable executives.

Automation, properly crafted and informed by policy, supporting good processes has the potential to reduce the risk of noncompliance by a significant amount, and keep senior executives out of jail. Ensuring that processes are documented and repeatable and providing timely and accurate reports to compliance officers are good defenses and may even be a real get out of jail card. There are other laws and regulations concerning data and privacy, such as Health Insurance Portability and Accountability Act of 1996 (HIPAA) that will benefit from the consistency and accuracy of automated reporting.

We have developed Table 3-1 to represent levels of automation and collaboration to put the tools and techniques we are discussing into context and show their relationship to other levels of automation. Level zero in this table represents the lowest level of automation and collaboration. The second level of the table refers to human-directed automation. A fixed logical process in row 1 refers to transactional and rule-based processes using structured data that acts as a precursor to RPA. Structured data in this context is data where the data structure is known and adhered to.

Table 3-1. *Levels of Automation and Collaboration*

Level	Human-Machine Interaction	Intelligent Automation
0. *Non-intelligent tools*	*No intelligent automation* • Non-intelligent devices that operate with little or no human supervision. • Examples: gas engines, boilers, water turbines that existed prior to 1930s.	*No intelligent interaction* • Human makes all decisions, and interpretation is fixed. • Examples: driving or braking in cars made before the 1970s; mechanical looms.
1. *Human-directed interactive tools*	*Human-directed automation* • Fixed logical process designed and initiated by humans. • Examples: batch processing and RPA; industrial robots working in restricted areas; the Jacquard loom with punch cards (c. 1801).	*Human-directed interaction* • Human makes all decisions; machine can make local adjustments. • Examples: modern antilock braking systems and cruise control in cars; standard text (w/ autocorrect) or graphic editors.
2. *Partial or conditional collaboration*	*Human-assisted automation* • Human selects goals; robot recommends actions and once confirmed acts with limited autonomy. • Examples: smart buildings automatically adjusting lighting and airflow; intelligent process automation (IPA).	*Machine-assisted interaction* • Human selects goals receives continuous feedback, and can quickly assume full control. Robot has limited autonomy. • Examples: virtual assistants that reserve flights for air travel; traffic-aware cruise control; remote-controlled surgical robots.

With these precursors in place, it is possible to use tools that require no programming to develop a resulting automation. Most relational database developers had tools that would allow users to create complex queries using a visual interface and a mouse. Programming resource, even for small tasks, is expensive and may not be released in a timely manner to fulfill requirements. Screen-scraping tools grew out of the need to alter and automate a task workflow without programming resource. Screen-scraping tools can be trained to follow a set of steps by capturing mouse positions, screen data input fields, and pressed function keys. A screen scraper can be instructed to copy data from an application, paste it into a data field in another application, use that data to search a database, and display the result. Once the screen-scraping expert has set up the tool, it gives the appearance of a new screen with only a few actions; all the underlying work of copy and paste are not seen. The background process and data stays the same unless it is the subject of a screen scrape operation.
There is a similarity between an expert creating a new look, task-based application on top of existing applications and the learning phase of an implementation of RPA to the extent that some industry analysts have commented that RPA is screen scraping on steroids.

Increasing the Automation of Business Processes

Automation of business processes to improve repeatability, throughput, and accuracy has been a goal for many years. Automation can be split into several domains as referred to in Table 3-2. This chapter is concerned with the first of these domains, business process automation.

Table 3-2. *Automation Domains*

Automation Domain	**Description**
Business process automation (BPA)	Also known as business automation, this is the technology-enabled automation of complex business processes. BPA can streamline a business for simplicity, facilitate digital transformation, increase service quality, improve service delivery, or contain costs.
Enterprise workflow automation	Looks at the hundreds of processes involved that keep large organizations moving forward. Enterprise workflow management identifies the best ways to map, execute, integrate, improve, and automate workflows. This is an operational tool not a business tool focused on ensuring that all the architectural elements in a workflow are available and optimized, for example, ensuring that there is enough disk space for a file copy operation.
Automated manufacturing	Integrates software and machinery so that manufacturing processes are run autonomously through computer programming.

In the age of IT, there have been many attempts to provide solutions to changing workflows of existing applications without changing the software. Some years ago one of the authors came across a good example of this in a print works. A print job was supposed to have a unique number and a customer reference. If the job was small, one of a kind, and from a new customer, the clerk is supposed to create a new customer record but often clerks they don't for a single job. They jump to validating the job, sending the job to the shop floor, and completing the paper work. They do this because the administrative tasks of creating a new customer record would take longer than the associated printing job, so the clerks used a dummy customer number 99999 for all such jobs. This had an associated customer ID of "Miscellaneous."

In a small firm this type of hack was not really a problem. Larger organizations could be vulnerable to fraud or fraud accusations from auditors, and in both cases, incomplete records could remove an opportunity to upsell to a new customer. This practice subverted a part of the business process and became a new process that removed the "enter customer details" task from the original process. These problems may be the result of poor analysis, poor workflow design, inadequate procurement requirements, intentional fraud, or poor application specification or weak regulation.

There have been many methods for automating work using IT. Figure 3-1 shows the evolution of automation tools over time and their relationship to automation tools. For example, screen scraping can be seen in this figure as a potential precursor to RPA. It also runs in parallel with later tools. Rule-based applications and screen-scraped applications are seldom discarded but the new tools are used for new problems. Intelligent automation in Figure 3-1 is the end goal of automation in this context. It can take the output of tools that are task based such as business rules engines or screen-scraped applications as well as process-based tools like RPA and combine them with AI to build a business decision tool. Automation tools such as RPA support business processes but do not change them on its own. Automation tools are not the silver bullet that can repair a process error. There is a risk that a process error will be executed faster by RPA and potentially create more problems.

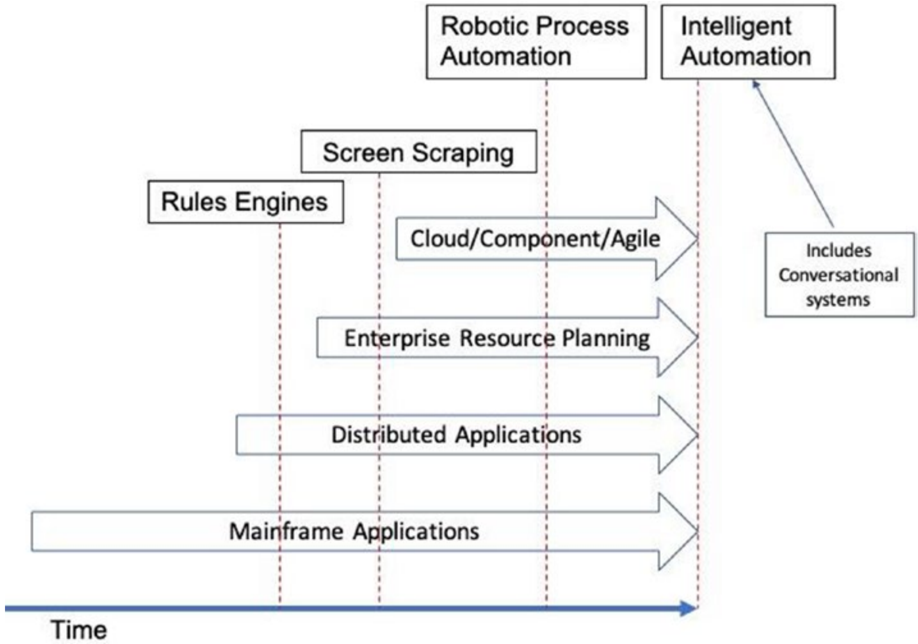

Figure 3-1. *Context of RPA*

A new automated application often has to coexist with other applications supporting different workflows. The new automated application may use the same components, services, and architecture. All of these different applications, old and new, are using different tools. Existing applications are frequently called legacy applications. Legacy applications are not just the mainframe applications noted in Figure 3-1 but can also be distributed or enterprise resource planning (ERP) applications.

Legacy application is a term that refers to an application that has been deployed for a long time and is frequently difficult to modify, update, or even maintain. Rules engines and screen-scraping tools first came on the scene in response to the high cost of modifying and updating mainframe applications. Rules engines were designed to add changes to inflexible applications by permitting parameters and decisions to be taken outside the application itself. Screen scraping can revitalize the GUI of legacy applications, converting green screen mainframe applications to GUI applications. These are both part of a move to increase the scope of automation and improve the speed of reaction to business process changes. In Figure 3-1 the timeline refers to the length of time that particular solution applications have been implemented with many mainframe applications being implemented in the 1960s and 1970s. Rules engines have been popular in the domain of mainframe and distributed applications, and screen scraping cuts across both domains and included enterprise resource planning. RPA is a process not a task-based tool so it also includes cloud computing and agile computing in its scope. The influence of the different tools on automation of business processes increases as you move from mainframe and rules engines through to screen scraping, but the automation of difficult processes with many components requires tools like RPA.

Cloud computing and continuous improvement enables the automation of business processes by making changes to the composition of applications. Using components or cloud services that are composed quickly with little change is a fast way of building applications with little programming resources, and the use of composition tools rather than application development tools gives more flexibility for changing workflows at a code level over the last 5 years than was feasible for older legacy applications. Some of these composition tools use a graphical or mapping language to enable point-and-click composition, again with no coding needed. Even with these tools there is a great deal of work in the location and selection of services to compost into a new application that supports the whole process. RPA takes existing application and task workflows, links them with other tasks and workflows across the whole process, and generates an automation from start to end.

Replacing one cloud service with another that better matches the requirements of the changing process enables rapid changes to the logic and workflow of business-supporting applications with a minimal coding effort. Automation using reusable services will use fewer coding resources than rejuvenation of legacy and possible monolithic applications but requires a different infrastructure from those monolithic applications. Often this architecture requires a large infrastructure investment. RPA needs no extensive infrastructure changes.

Agile methods and continuous delivery tools will deliver applications that can react to changing business requirements that can require weekly or even daily updates, but delivery of these applications will use considerable resources. Neither of these automation strategies and the multiple variants of the strategies can deliver with no coding. RPA can have a quick and inexpensive impact on business by working at the user interface level with no additional coding. We will discuss the RPA contribution to increased automation later in this chapter. First, we can consider here the impact of earlier automation strategies on business.

If business automation needs to change to support changes in the underlying business process, we should consider the impact that changes to the business process will have on business automation. There are very few businesses that have no competition and do not have a need to modify their business plan or software. Some years ago during a roundtable discussion with CIOs, CFOs, and academics, one of the topics we considered was the potential for survival of businesses that have little or no change to their business model. The roundtable was convened by CA Technologies to plan a research project into business model value. An example from one of the CFOs was of a factory he knew that produced gold thread for military uniforms and braid. This was a very niche market that at one time only two companies worldwide supplied all the military gold thread and braid. These two businesses still provide all the high-end shops and military suppliers but their niche market is being encroached upon. Checking recently, it seems that there are now more factories involved in this business that seem to be branching out into a number of other areas. This will bring pressure to change business process supporting applications. Even niche markets change over time and may require software automation to remain competitive and survive.

There are clear advantages in large enterprises making changes to the logic of underlying applications or using replaceable services to generate adaptable logic that supports new business model changes. These changes will require design, development, coding skills, testing, and release processes that are often outside the capacity of small- to medium-sized companies that nonetheless need to react to business model changes.

Many organizations have adopted and implemented agile, continuous improvement methodologies to facilitate service development and usage, but the cost of creating an IT department capable of managing this process is frequently seen as high by many companies. The cost of entry to agile development methodologies may well be more than a small company can absorb, and even large companies have to consider the cost of agile as part of their improvement and automation strategies.[2] RPA has a much lower cost of entry that makes it more suitable for these small- to medium-sized organizations.

Process Management, Selection, and Optimization

RPA does seem to polarize opinions. Analysts and reporters in favor of RPA state that RPA is the next step in intelligent automation and is a precursor to full AI-led business process management. Other commentators, less impressed, suggest that RPA is merely screen scraping on steroids. Even though it is scalable and has extra functionality, they point out that screen scraping never really worked. Since our discussion is about the impact of RPA on automating processes, we will leave the debate about the effectiveness of RPA to others. We will also confine ourselves to discussing

[2]Embracing Agile; D.K. Kirby, J Sutherland, H. Takeuchi, Harvard Business review https://hbr.org/2016/05/embracing-agile

the broad architectural functionality of RPA rather than the products and technologies that focus on this potential market.

As mentioned earlier, increasing automation of business processes has been a goal for many years and one of the technologies that showed initial promise was screen scraping. Screen scraping used the position and contents of fields on a screen to emulate a user of those screens by copying data from one data field on the screen to another or using the data to develop a query in the background. Screen scraping can also take a difficult navigation through an ERP system and simplify it to a single key press. Screen scraping enabled new screen design at the user interface level and underlying logic that could change the way that the host applications were used.

At its best, it has rejuvenated mainframe, green screen applications, making them easier to use for the infrequent or novice user, presenting actions and results in a relatively modern interface. Screen-scraping tools produced user interfaces that were never intended to replace the expert user but to support the infrequent user enabling them to navigate to a functional screen, for example, navigating to expenses approval screen without using a manual and six different command keys and screens. Expert users of a green screen application are faster than a novice user, but novice users frequently object to the "waste of time" using those green screen applications.

In some screen-scraping implementations, it was even possible to use a mobile phone to carry out parts of a mainframe-supported business process. There were a number of problems with screen-scraping-based user interfaces, but the most critical was the requirement to change the screen-scraped application if the underlying screen fields or command codes changed. It was not adaptable to even trivial changes to a mainframe green screen layout.

The next most critical issues were performance and scalability. Scalability often became an issue relating to the performance of underlying applications and hardware impacted by numbers of users of the core systems, the interrelationship between different core systems suppling the data, and data management issues. Performance is similarly dependent on the performance of the core systems. It is claimed that RPA does not have all the same limitations and also has additional functionality that makes it more attractive to users than screen scraping.

Screen scraping is an action- and task-based strategy using base applications that must be built by point-and-click type operations, while RPA offers the opportunity to combine actions and tasks performed on multiple applications into a seemingly new application, built by copying the actions of a user following tasks in a business process rather than tasks dictated by a core application. RPA evolved from virtualized testing tools which recorded a user's action and played it back to automate tests. RPA uses the similar concepts of creating test scripts from application usage to create automation rather than testing.

The enthusiasm for creating automated processes in a simple way, by replaying the usage of applications, can cause security to become overlooked. A software robot, developed as part of an RPA implementation, can be treated as another employee, having the same security constraints, authenticated and granted access to systems to enable it to work. This would require a sophisticated authentication and access control policy that is beyond the means of small- to medium-sized enterprises (SMEs). In many ways the security attack surface is no greater than an application that is in common use and the issues are well known.

RPA has some specific security vulnerabilities due in part to the architecture of an RPA solution. A human may see an instruction or an operation and say "this is odd" an RPA robot would not ask why. There are also concerns that a badly trained software robot may intentionally or accidentally violate compliance rules. Other security breaches may come through the development route.

In a normal development cycle, there are design and coding rules, oversite and review processes, and testing that all may validate the modifications. Even in the agile methodology, daily stand-up meetings can expose errors. RPA being trained and in effect built by users demonstrating the business process can bypass oversite activities bypassing some user privileges. It is also important that suppliers, developers, and trainers of RPA implementations are transparent and accountable and that the solution is QA tested including security testing.

All software can be classed as a robot and some software robots are smarter than others. In RPA the interaction between applications is handled at the user interface level. The data is captured from an application in the user GUI layer and passed to other applications at that level. It does not need a framework to pass data between APIs and it needs no change to the underlying architecture.

To develop a software robot, the RPA toolset observes a process, records it, and then plays it back as though using a screen, in the same way that virtualized test data and interactions are recorded for testing. RPA uses the actions of an expert user for training. It relies heavily on earlier incarnations of virtualized testing. RPA does not need hardware modifications and that removes the complex management issues around new infrastructure and applications. Some RPA instances require human assistance to work and are referred to as assisted RPA.

Some RPA instances are fully automated. Fully automated RPA needs no human intervention and supplies the type of activity that will have the most impact on work in the future as it is more likely to replace humans, but all RPA currently operates at the task level not necessarily the complete, end-to-end process. RPA changes how services can be delivered by replacing people with technology. We are not talking of technology enablement where desktop scripts assist human agents but software automation that replaces most or all of the work previously performed by people. There is a heavy human price to pay for some automation, and this style of automation is happening now and for the next 2 years is expected to increase the displacement and possible dismissal of knowledge workers.[3]

RPA Implementations

After a review of the potential of RPA, most organizations' next steps are to decide if, where, and how to implement RPA. This is usually driven by a need to reduce costs or improve services. There is often a feeling or even a hope that the process itself is likely to improve.

As we will show later, this is a false understanding of the impact of RPA. RPA does not change the business process; as the previous section notes, it is nonintrusive and works at the user interface of the supporting IT processes. It can improve the speed and consistency of a process, but it cannot alter the process at all. If there are flaws in the underlying data or in the process flow, they are still there after the RPA implementation.

Implementing RPA focuses on a business process. The business process is unlikely to be automated; as a whole there are often tasks that require human intervention, such as approvals, not eligible for automation. In Figure 3-2 you can see a notional set of tasks forming part of a business process.

[3]Willcocks, L., & Lacity, M. (n.d.). Nine likely scenarios arising from the growing use of robots. Retrieved May 16, 2019, from http://eprints.lse.ac.uk/64032/1/blogs.lse.ac.uk-Nine%20likely%20scenarios%20arising%20from%20the%20growing%20use%20of%20robots.pdf

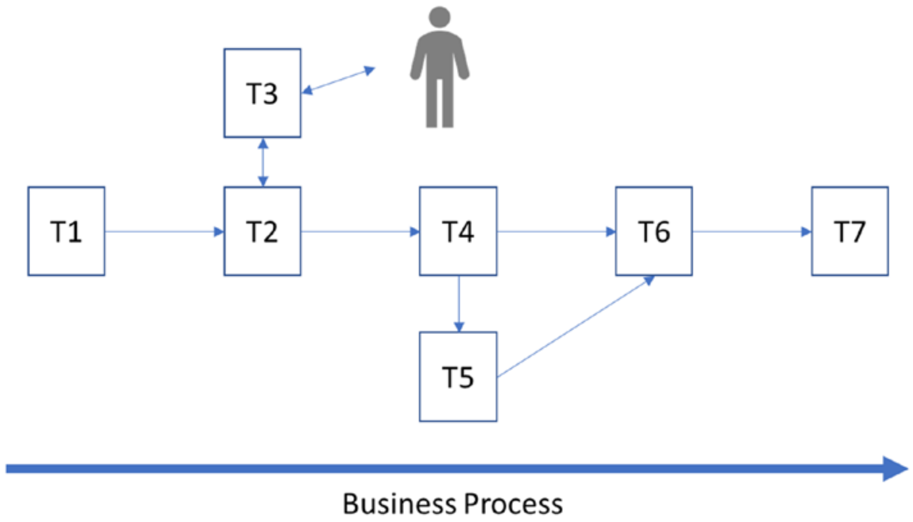

Figure 3-2. *Business process and task relationship*

In this business process the process flow is between tasks, some in sequence, some depending on human intervention, and some that can be done in parallel. Human intervention cannot be automated, but each task can be automated using RPA if they have a defined set of repeatable rules. Cloud services, when first introduced, allowed line-of-business departments of a company, like marketing or sales, to use online services without contacting the IT department. A sales department could decide to use Salesforce as a management and monitoring tool for their sales processes without contacting the company IT department to handle procurement and adoption of a sales management tool. This became known as shadow IT. Thanks to its low cost of entry and relatively simple installation and operation, RPA has been regarded as similar to shadow IT. Implementing technology outside the IT umbrella has a number of potential issues but is usually a tactic to implement quickly and avoid the potential stigma of being an IT implementation. Departments considering implementing RPA outside the IT umbrella should think more strategically.

Shadow IT is tactical and IT needs to be brought into the project as an article by InformationWeek noted.[4]

RPA can also be used in organizations where the IT department may be small. RPA and shadow IT generally become a problem if the IT department does not know about the use of RPA tools in the company and cannot evaluate the risk of violating the various data protecting and privacy protecting regulations. An additional issue would be the potential for unanticipated budget and resource shortfalls in the IT department should the department using RPA decides that it is no longer prepared to devote their own resources to managing RPA and "throws it over the wall" to the IT department. This has been an issue in shadow IT and has the potential to be an issue in RPA implementations.

Once an opportunity for RPA has been discovered, the first step is to study the business requirements and select the appropriate tools. Before making any decision on technology, the process to be automated must be selected. Process selection is critical to success and can help resolve a number of issues, but before selecting a business process, it is important to review the existing processes. As you can see from Figure 3-3, simple, repetitive processes are the low-hanging fruit of process selection and will drive value from early RPA implementation. As part of the selection, processes need to be evaluated in terms of the operating environment in which they exist.

[4]Morgan, L. (2020, February 18). Who Should Own RPA? - InformationWeek. Retrieved April 6, 2020, from www.informationweek.com/big-data/ai-machine-learning/who-should-own-rpa/a/d-id/1337065

	Manual	Cognitive
Easy to automate	Simple Repetitive Can be Anticipated e.g. Data Entry	Complex Mostly Repetitive Can be Anticipated e.g. Insurance Application
More difficult to automate	Complex Creative Un-Anticipated e.g. Premium negotiation	Complex Creative Un-Anticipated e.g. Evolving Process

Figure 3-3. *Automatable task classification*

Figure 3-3 illustrates the way in which tasks can be classified. Simple manual tasks, such as data entry or onboarding new customers and suppliers, are likely to be automated. Onboarding new customers, for example, may rely on emails and web orders. This process may have a manual component if the email or online form is not complete and may only be complete when a human makes further enquiries for any missing information. The manual element of the process would then be less focused on the data entry and data validation and more focused on creating that additional information. Cognitive processes that use a formalized set of rules to complete tasks, for example, calculating an insurance premium, can be easily automated. If the decision is challenged or the customer wishes to negotiate with a clerk, this would fall under the complex banner and be more difficult to automate, although chatbots may change this.

For example, the process to create a new customer order may contain a number of steps that can be automated. The process may have different flows depending on the operators. Experienced operators may have a slightly different method, for example, a particular stock item may be frequently ordered making its stock number memorable. An experienced operator may merely type in the relevant stock number instead of searching for the stock number. When automating a process, the operator being used to train the RPA tool should use the basic process and not any shortcuts that they have developed over time.

Selection of a process that would benefit from automation is the most critical part of implementing RPA. Often the overall success of RPA in an organization depends on the automation of well-selected processes. Selection criteria for processes to be automated may vary but broad guidelines include the following:

- Pick processes that have repetitive, high-volume work that requires consistent approaches. Over time tiredness and boredom can reduce human efficiency and accuracy.
- Ensure that the process is well understood, and that allowance is made for any shortcuts or local knowledge.
- Establish clear objectives that are well understood. Is the RPA exercise part of a general strategic business review or focused on a specific tactical goal such as improving speed of throughput, reducing customer wait time or frustration?
- Select a currently stable process that is not scheduled to be modified or included in a new business process.

Resistance to change, particularly change supported by automation, is a factor that can't be ignored. Automation as part of cost-saving exercises can make staff nervous and lead to worries over employment or wholesale firing of staff. Good change management is needed to counter this and RPA implementers need to consider the human factors as much as the technology. An organization that is ensuring that their employees see a clear migration for them from repetitive and dull work to work that is more worthy of them will manage the transition more successfully than organizations that ignore these important factors.

Earlier in this section we mentioned the excitement and rush to implement new technology solutions and gain benefits quickly, but this has to be tempered by a need to ensure that the best processes are being automated. An excellent implementation of automation for a process that has flaws is no better than a poor implementation of an optimized process. Business process management offers methods to review and optimize processes, but this may be outside the scope of the small- to medium-sized businesses where RPA could deliver advantage. There are a number of simple questions that can be used in the review of processes and their associated tasks to decide if they can be automatable:

- Are there any unnecessary steps in the process?
- Are there any redundant operations?
- What is the impact on the process of removing unnecessary or redundant steps?
- Can the process be simplified in any way?
- What business rules are codified in the process?

While this is not an exhaustive list, it indicates that there are a number of challenges to implementing RPA beyond training the automation.

RPA: Advantages, Challenges, and Caveats

The advantages of implementing RPA have been outlined before, but it is worth outlining them again in relation to challenges that RPA presents. During late 2018 and early 2019, RPA became more readily accepted as having an impact on businesses. These are businesses who are taking the opportunity to revitalize older IT and business processes, particularly manual and repetitive tasks with a high error rate. Cost savings and productivity are the main drivers for implementing RPA with an ROI of 5:1 claimed by some organizations.[5]

The architectural style of RPA implementations, with the software robots sitting on top of existing IT systems and able to release staff for other work, is relatively easy to deploy and gain early advantage.

- RPA will generate lower error rates for many tasks, as it does not get tired or bored.
- RPA can easily harvest data from emails, online invoices, or payments and put into ERP or CRM systems.
- The ability of an automated process flow to manage data from a variety of different systems can enhance performance and contribute to increased regulatory compliance.
- Regulatory compliance is enhanced because rules that support compliance are applied at the same stage in a task every time.
- RPA tools can deliver improved analytics and reporting as well as employee satisfaction.

RPA does not replace all human functions in a set of tasks, in a process flow such as that in Figure 3-2. Some tasks can be automated and others either cannot be automated or require some form of external validation. This could be for validation or additional information and there may be judgment calls that have to be made on exception.

There are some "low-hanging fruit" in terms of automatable processes. In Figure 3-3 the tasks that are repeatable and simple are all good candidates for automation, and while they can yield good ROI, the

[5]RGP, https://chapters.theiia.org/san-diego/Documents/Presentations/RPA%20IIA%20Presentation%20San%20Diego.pdf [accessed on August 2019]

more difficult tasks are ones that may give greater benefits. It is more difficult to automate customer service tasks that require human-to-human communication, but if the supplier side of the communication can be automated using chatbots, then not only will this free up staff for more personalized discussions or more challenging tasks, it could free up an organization who can then offer 24x7 support without the need for staff in different locations or shift work.

Even some of the more abstract decisions taken by a human can be codified if there are some basic rules. If a decision-maker such as a hiring manager is asked how they make a decision, they can describe a number of rules that they apply. They may look to cross-referencing candidate's qualifications with desirable qualifications, for example. If the decision-maker also looks at length of time in previous company ideally > 1 year, then this can also be applied to all the résumés that are presented. There are underlying rules that may not be codified, based on the manager's experience or personal knowledge that could be used to further refine the search. This experiential understanding can then be addressed and potentially codified.

Staff will happily hand over tedious or boring tasks to RPA and may also appreciate an improvement in accuracy. More difficult to automate tasks will cause more potential unease since many of these tasks may be the only expertise of a member of staff; thus, they will feel more threatened. Poor change management and communications from management can leave some valuable staff concerned. This should be addressed in the implementation stage, and one of the major challenges in implementing automation is in the area of staff relations. Not long ago, one of the authors met with some customers who were planning on implementing payroll automation software, and I said, "This will speed up and simplify the work of a payroll clerk so that they can spend less than a day preparing and issuing pay slips." He did not realize that one of the people in the meeting was a payroll clerk and their sole job week in week out was preparing the payroll and thus they were very threatened.

Any changes to business models resulting from RPA need to be communicated and the redeployment and retraining for the displaced staff should also be considered from a business model point of view. Business model changes that place less emphasis on staff and more on automation can increase the threatening nature of automation.

An additional challenge is the management of the implementation. This may seem self-evident, but as was noted before, there are issues regarding change management and communication that can derail implementations. There are a number of high-level decisions and tasks that have to be acted on at the start and during the project, for example:

- Identification of processes for automation.
- Process optimization, reviewing current processes, and see if there are unnecessary steps.
- Managing change and relationships between IT and LOB.
- Who runs the project?
- Continuous improvement, retrospectives, inspect, adapt Re: Draft Local Management Agreement.
- Managing staff whose job is being replaced.
    - Redeploy.
    - Exception review and audit.
    - Extend role with more intuitive/creative work and more robots.

Although the robot can perform faster than a human, the robot can only proceed as fast as the process will allow. For example, if a correlation search between two databases takes 1 minute, this part of the process will take 1 minute whether the robot has initiated the search or a human operator. Software robots must obey architectural rules and compliance for the base system.

If as mentioned earlier there is the possibility of an ROI of 5:1, then any outsourcing proposition would have to be based by balancing the risk of automating a task rather against sending repetitive tasks offshore. Risks can include cost and ROI, speed of automation against training offshore staff, and normal offshore challenges of time zones, languages, and quality. Chatbots will also affect offshoring. The risk to offshoring is based on conversational systems using core applications as automated help desks and call centers. Chatbots have this potential, but RPA does not affect offshoring. Some commentators see offshoring increasing rather than decreasing. Any existing offshore hosting will probably stay offshore, and a little confusingly reshoring will also grow.[6] Reshoring and insourcing of processes and applications is the methodology for bringing work back into the country or the organization.

Another issue that will not magically go away when you apply RPA is the relationship between line-of-business management who own the processes and the IT team who develop the IT business–supporting software. The disconnect between these two teams is often great and RPA is not going to change that, unless it finally becomes the technology that can allow LOB to develop their own systems.

Intelligent Automation, Bots, and Chatbots

Through this chapter we have been talking about RPA implementations that have to be taught; learning and recording mouse moves as the task is accomplished by an expert. When you look into the future of RPA and automation in general, it is easy to fall into the trap that AI will solve everything and enable software robots to automate all but the most difficult tasks in a process and automate whole processes as well.

[6]Nine likely scenarios arising from the growing use of robots. L. P. Willcocks, M. C. Lacity, Retrieved May 16, 2019, from http://eprints.lse.ac.uk/64032/1/blogs.lse.ac.uk-Nine%20likely%20scenarios%20arising%20from%20the%20growing%20use%20of%20robots.pdf

AIMDek makes the point that there are several types of RPA, including assisted RPA, where a human is supervising the RPA activities, and unassisted RPA where the software runs on a server without human intervention,[7] although the process may call on other servers and applications. Figure 3-2 showed that this process is made of many tasks. Some are done in parallel and some depend on other tasks completing before they can act. Unassisted RPA can automate a task that needs no human intervention, for example, an unassisted RPA may complete all the tasks to complete a mortgage application although the process may need an intervention, final authorization from a human to complete.

In the world of software suppliers, there are conflicting views of the future of RPA and a lot of terms that are being used to try and identify each supplier's view of the next step in process automation. The recent consensus is that RPA as we know it will evolve into cognitive RPA (CRPA) or cognitive automation (CA)—cognitive RPA (CRPA) where AI can be leveraged using technologies such as optical character reading. Cognitive automation is pretrained to automate specific business processes. Both of these are similar in scope, anticipating a use of AI to make process automation stretch ever further into complex tasks and processes. Bringing intelligence to bear on unstructured data as well as structured data will be the next stage on from the RPA that is currently being implemented.

EXPERT INTERVIEW WITH SERGE MANKOVSKI

Serge Mankovski is an expert in workflow automation and holds a large number of patents in this area and he also has some interesting things to say about intelligent automation. In this interview, we discussed his views on RPA and intelligent automation and based on the work that he has been doing in

[7] AIMDek Technologies, Evolution of Robotic Process Automation (RPA): The Path to Cognitive RPA, August 29, 2018.

workflow automation, AI, and machine learning. We have made some changes for readability.

- Intelligent automation as a term appeared a few years ago. The earliest references were to enterprise automation and workload automation. Many of the tools in this domain had AI components but these were not a central feature. Critical path analysis had some AI tools and intelligent routing also had some intelligence, but it was not until chatbots and robots became mainstream thinking that everything became an intelligent automation system. There is an element of old wine in new bottles but there are some systems that are becoming more intelligent. A good example would be a document management system, which is in effect a gigantic workflow system. There are some AI-type components to make some intelligent routing decisions although again they are not central features. The advent of chatbots and robots creating the hype around intelligent automation showed that there are still the same issues that must be solved.
- The concept of automation has expanded over time, but it is still not well established. There is little open source, few standards that address this domain, and continuous delivery has generated islands of established system that exist in the world of agile and continuous delivery.
- Intelligent automation took enterprise automation by surprise. Enterprise automation was used to build large system environments with well set-out procedures, building very large systems. These systems had version control and other processes to automate the management of these very large systems. Agile computing plus the process and tools that support it have a different infrastructure and also enable a mobile infrastructure. This is a new type of automation based on agile processes and

infrastructure. This type of automation will use hourly or daily delivery of updates. It would use daily integration within the delivery window using tools such as Docker and Kubernetes and be influenced by open source. As was said earlier, the preeminence of open source has blindsided many tools and infrastructure providers. They are struggling to monetize their products and still compete, how to charge for company products. There are lots of open source tools that don't get charged for; agile tools, for example, are generally free. Commercial agile tool providers have to rely on the attitude of customers who want to offload the responsibility for managing open source tools by using corporate branded tools.

- There are a number of islands of automation in enterprise IT, document processing, and intelligent manufacturing. Intelligent manufacturing, for example, is well advanced with solutions in logistics and supply chain that are dominated by niche players. Automation is a tapestry with many minor components like data preprocessing pipelines, data pipelines that are daily requirements and need to be repeatable and at scale. There are requests for new tools, but they are not being fulfilled completely. The focus is on intelligent automation rather than creating new tools.
- Intelligent automation would consist of several expanded layers, not just a sole emphasis on chatbots and "shiny things" that still need an infrastructure around them. Organizations are by and large still building the old-style proprietary infrastructure. Automation needs are now different and so are the solutions. Some companies are providing tools that bridge the islands of automation, stitching together the different islands but still not sure what intelligent automation is and what are the components.

- One thing that is clear is that chatbots and other shiny things need to be integrated with enterprise automation. This can then be supported by an infrastructure that tries to weld together proprietary solutions. This could then give the possibility of updating software in robots through a continuous delivery pipeline.
- Disparate robots can communicate, using topic-based pub/sub in the Robotic Operating Systems (ROS). As an open source meta-operating system for robots, ROS will ultimately enable high-level commands to be interpreted and become actions. An example would be the entry of a robot into an intelligent building. It would require a sophisticated way of coordinating with the buildings. There may be goal conflicts between the robot and the building, perhaps to deny or permit entry based on the robot's intent. There is a need to be able to do this at the semantic level, with the semantics as simple as you would use an API-level communication.
- If you assume a well-structured robust API, that is well tested. There is a need to be able to say at the semantic level, "What have I accomplished?" and "What is the high-level goal?" It can be important to play this through a scenario, for example, using these questions:
    - What will happen if I accomplish this goal?
    - Am I going to fail?
    - What additional resources are needed?
    - What kind of capacity?
    - Does this violate any constraints?

- These are all intelligent automation components. With these being managed, we can look at a system and it can do things sensibly, that is, reliably understand what the system is doing and how to progress. If you look at the question about failure, it is possible to evaluate the risk of a failure and the subsequent cost. How can you reconcile intelligent automation with enterprise workflow automation, the automation of IT tasks and business tasks? If the cost of failure is high and the risk (or likelihood) of failure is low, then there is no issue with failure. There is nothing in intelligent automation systems at the moment that can make reasoning to recover from a failure automatically. Intelligent automation has to grow pragmatically and should cover the mundane not the shiny.
- One direction is to build on the blackboard problem used in problem escalation. There are many AI tools on a message pipeline with intelligent parsing of messages that throw any estimated problems on a "blackboard" for review. The review is managed by humans who can resolve the specific problem event and the system learns from each review so that they get better at recognizing problems and giving fact-based answers.
- The next generation of sensible AI has humans in control and works synergistically with AI. In this model the machines mine data and propose results with the humans making decisions. It's not a new idea at all, but it needs work to become a scalable, repeatable, and sensible workflow. Intelligent automation will get human performance on the critical path, but despite some commentators, it will not reduce staff but increase the scope of the existing staff. This will also require an investment in continuous learning and staff development, for example, data scientists may have to take note of and prove their expertise in all the new "stuff" that comes out over a year to maintain their certification.

- The good news is that the critical mass is already here. Technologies exist that cover a lot of the AI domain. We already have the technology that can build large parts of IA solutions. A good example is Google deploying tensor flow and showing the world how to build large-scale ML models. The problem is resources; you need to put really smart people on it and professors are already saying that they can't find good PhD candidates in Silicon Valley because the students are going into companies to do ML. Cloud providers are sucking up data; ML can point to this data and use it for reasoning, although the cloud providers will charge for access to the data. To model this, we need a good use case, not autonomous cars, but healthcare may produce an interesting result.
- A critical element is the theorem of universal approximation, currently handling fake news, movies, and so on. This can conduct reinforcement learning and can be taught to automate any process as long as there is a good fitness function. It can replay continuously on its own and develop rules through adversarial learning as well.
- There are a number of gaming companies who create world simulation and a good 3D simulation, from Unity, for example, can produce a great simile of the world for low-risk experiments. If you take a population of robots that can learn in a simulated environment, then you don't need to build a physical automated system to evaluate intelligent automation. This would be a complete disruption to the current models. It would also be possible to create intelligent entities and throw them into a virtual world that behaves by the rules that you want to observe. If you want a rule that says stay on the red line, you need to design an appropriate fitness function. The robots will learn in the virtual world at little or no risk.

- Reinforcement learning aims to train an intelligent (automation) agent where as other forms of learning aim to make predictions or estimations. Reinforcement learning (RL) is there to teach machines to automate tasks by maximizing the reward so that the machine can exhibit emergent automation behaviors instead of explicitly prescribing it. This is in essence how robots learn to learn to walk nowadays. It is interesting that RL allows intelligent agents finding new automation strategies never discovered by humans. This phenomenon is somewhat superhuman, but it makes it more imperative to erect an intelligent cage around the intelligent automation so that it does not develop models that would violate policies and constraints within which it must operate.
- Creation of this intelligent cage is perhaps the most important aspect of intelligent automation of the future. It is akin to the need to develop deepfake detectors we are witnessing today. By fighting AI-based deepfakes with AI-based deepfake detectors, researchers pave the way for AI-based intelligent cages for intelligent automation agents. This cage is a firewall for the integration fabric joining islands of automation into a seamless flow defining the boundaries for behaviors of the process as a whole.
- Many of the current AI companies use single AI entities doing single-purpose tasks. Multiple agencies can act on the virtual world with no risk except compute time because it is all in the VR. The biggest problem is designing the fitness function and creating automation relies on the design of fitness functions.

- The future of AI will come when automation becomes optimization. We know how to do optimization so that we can reformulate from rules to fitness functions. Once this happens IA will explode. It is not a doom and gloom scenario; we need to create the rules of engagement. We need to become controllers managing machines rather than controlled people being managed by machines.
- Specifically regarding RPA, large companies can afford to develop RPA with optimized processes; they can hire logistics and operations people to optimize processes. SMEs don't have the same ability and some domains are better at this than others.

Some additional key points that Serge mentioned are as follows:

- Hype has generated interest in automation although the reality is a lot slower than we are led to believe.
- The interest and level of automation has expanded over time.
- There are few standards and open source solutions and this encourages vendor lock-in and makes it difficult for interaction between different solutions.
- There is a need to automate the underlying infrastructure to enable performance at scale and resilient infrastructure to support the RPA processes.
- Virtual trials of AI and robotic tools can be made using gaming platforms.
- Critical mass for AI is already here; there is already enough technology to cover most AI domains.
- As always resources are a problem, professors are saying that they can't find good enough PhD candidates for AI.

While a number of the points here are in line with our thinking, the use of 3D simulation to experimentally test theories of robotics is not quite as advanced as the comments make out; it is already a well-established model. Indeed, we spent many meetings in 2018 discussing the potential of using gaming models to test theories of robotic collaboration in a lower-cost, low-risk, environment without a physical engineering laboratory. It is particularly valuable in generating the large volumes of data needed for experimentation.

We have already noted that RPA does not change processes and this is supported in these notes from Serge. He also emphasized that processes need to be optimized before implementing RPA. It is clear from these notes and the rest of the chapter that there would not be a wholesale loss of employment but a migration of workers from repetitive jobs that are automated to more creative work that would be more interesting. The future could also offer facilities for small- to medium-sized enterprises that are currently the domain of large enterprises. While large-scale organizations can employ teams of process and IT optimizers, the SMEs can still automate thanks to the low cost of entry of RPA.

Summary and Conclusion

RPA technology is in use now and is assisting organizations to automate high-volume, high error rate repeatable processes with well-known rules. This low-hanging fruit can be automated by offshoring, but the number of tasks that need automation is vast and there is an appeal to automating in-house. It is often supposed that RPA, like so many other automation strategies, will result in large-scale staff losses. It is becoming accepted that staff will be displaced but not replaced by RPA. They may be deployed into jobs that are more worthwhile and less repetitive instead. There is, however, no doubt that RPA will be an employment disrupter for at least the next 5 years. Strategies for change by implementing RPA must take the human factor into account to more effectively develop automation solutions. Solid change management strategies will be key in avoiding the stress of imagined job losses.

In terms of the perception of adoption of RPA, there are mixed results, with some analysts suggesting that RPA is already at critical mass of deployment. This is explained by the low cost of entry and AI and robotics being at the early adopter stage. Process management and optimization are important techniques during an RPA project. This chapter focuses on RPA and its impact on business; that impact can be lessened if the corresponding techniques are not used to ensure that the automated project is in the best possible shape to deliver the maximum effect.

There are varying views of automation in general and RPA specifically. In terms of general automation, Chapter 2, "*Technology Definitions*," and subsequent chapters review opinions ranging from "the robots are taking over" to robots will "cause mass unemployment." The RPA predictions seem to be more pragmatic, perhaps because more people have seen real evidence of value. The value of RPA implementations seems to be rising and may increase more when small- to medium-sized organizations are looking for automation technologies with a good ROI. RPA will affect outsourcing by automating many tasks that could be outsourced; however, there are so many tasks and processes to be outsourced that outsourcing will grow in parallel.

An area of potential concern is RPA's potential to be deployed by line-of-business managers, in the style of shadow IT. One reason why line managers are keen to avoid the IT department is the perception of large impact IT project failures in a number of companies. There have been a number of large and small IT projects that have failed. The failures have been for a variety of reasons, massive budget overruns, poor functionality, poor performance, and so on. Line-of-business managers are often keen to control a project affecting their business. This can lead to problems of integration with existing systems, compliance with corporate standards, and increasing budget demands by the line of business. Line-of-business and IT implementers should work as a team to resolve these issues.

In conclusion, we find that:

- RPA will be a massive disrupter of business, work, and employment.
- The majority of employees will be displaced into other roles.
- Processes will need to be selected and optimized to gain the most advantage from automation.
- SMEs will find the low cost of entry attractive.
- In the future, employment will be disrupted further as more complicated tasks and processes will become automated by RPA.
- Intelligent automation using AI must have humans in control, on the critical path, and the automated tasks become more complex.
- Intelligent interactions between tasks, processes, and human interlocutors will be the next stage of automation and are discussed more fully in Chapter 4, "*Robots in Teams*."

Finally, this and subsequent chapters show how much closer to the desired result we can get: business managers finally being able to use natural language to describe a new process in a dialog with a bot that already knows or can easily learn your business. This will revolutionize business since the supporting applications will be supporting the business without the need to translate for IT.

CHAPTER 4

Robots in Teams

Devastation! If you were standing at the corner of Harbour Avenue and Front Street on this Caribbean Island, you would see in every direction, through patches of mist and dust, heart-wrenching scenes of devastation caused by the recent category 5 hurricane.[1] The houses and buildings of this once-quaint, popular tourist city are smashed, broken into timber and glass and twisted metal. Days later, when the winds subside to tolerable levels, fires still rage.

Throughout the city are the sounds of search-and-rescue drones. It would be dangerous for humans to stand at the corner of Harbour and Front, but robots of different sizes and shapes slowly move across the ruins. Drones search for signs of life using infrared and sound sensors. Land-based robots, based on extraterrestrial rovers designed by NASA, carefully crawl through the fragile and dangerous terrain. If the ground gives way and they tip over or fall, they can stabilize and crawl back to solid land. Hospital pods are airlifted to stable and central locations. All of this is accomplished using remote-controlled and semi-autonomous robots. They communicate with each other and with human medics and supervisors, forming small teams. Each team is responsible for a well-defined region of the city. The data they collect is used to map working and damaged infrastructure, and to trigger further surveillance, to predict

[1]The island and details are fictional.

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_4

safe and unsafe regions, and to automate various recovery tasks. In safer neighborhoods, survivors work alongside the human-machine first responder teams.

There are scientific challenges that must be overcome to create collaborative robots of the sort described in the preceding narrative, but much of what is described here is quickly becoming reality. Search-and-rescue is an excellent example of how robots can be used to complement or act in place of humans. Many of the tasks are dangerous and unsanitary. They require continuous vigilance, and extra-human strength and senses.

An Introduction to Cobots

In this chapter, we focus on collaborative robots, or *cobots*. They have also been called social robots. We use a broad definition—cobots are robots that work with and alongside human workers, assisting them and collaborating with them to manipulate physical and logical objects (i.e., data) in order to achieve an objective. Throughout this chapter we will use the term cobots to distinguish robots that work alongside humans from the industrial robots that work in separate spaces, usually for safety and efficiency.

The fundamental requirements for industrial robots (the kind that don't typically interact with humans) are the following:

1. Integrating (or fusing) data from various sources—information derived from sensors and other computers, and transmitted through communication networks.

2. Performing actions within its physical and logical environment in order to affect some well-defined objective. This often requires sophisticated

> decision-making software that accommodates changes in the data and instructions received through its internal and external communication networks.[2]

Cobots must fulfill these two basic requirements but must also be able to carry out their actions within a social environment. As in the search-and-rescue example in the chapter introduction, the cobot must be aware of the humans within its environment, evaluate the role of these humans (emergency worker, volunteer, patient, etc.), and appropriately communicate with and act alongside them. In the next section, we'll explore how cobots are changing organized work.

Cobots in Complex Environments

> *The gadget-minded people often have the illusion that a highly automatized world will make smaller claims on human ingenuity than does the present one … This is palpably false.*[3]
>
> —Wiener, N. (1964)

Cobots are designed to work alongside humans in complex tasks. This contrasts with traditional industrial robots that are intended to operate in a physical space that is separate from human work areas.

Traditional industrial robots have been used for decades in high-volume, high-speed applications (e.g., sorting mail, welding, or injection molding where robots can be isolated from human workers). In these types of applications, the robots need to achieve highly accurate

[2]The mapping from data fusion to intentional action is often mediated through a model of the environment, which may be learned or predefined. Alternatively, the mapping between data and action may be handled by a human operator who monitors and issues commands.

[3]Wiener, N. (1964). God and Golem, Inc: a comment on certain points where cybernetics impinges on religion (Vol. 42). MIT Press. p. 63.

movements, work with other machines and robots, and operate with little or no supervision on tasks that automate processes or products that don't frequently change.[4] Industrial robots are designed to accomplish highly routine, precision tasks, at high speeds and low cost—they replace workers or accomplish tasks that are too dangerous, too tedious, or too difficult for humans.[5] They are essentially the physical embodiments of the automation discussed in Chapter 3, "Robotic Process Automation."

RPA and intelligent automation facilitate the integration of other industrial processes (management and control) with physical production, including those that use industrial robots. According to Fortune Business Insights, the rapid shift from manual labor to automation is creating a rising demand for industrial robots. The global market size for industrial robots was valued at almost $19 billion (USD) for 2018 and is expected to achieve almost $60 billion by 2026.[6]

Cobots are a different species of robots. Collaborative robots are designed to work with or be trained by humans. They are lighter, safer (for humans), more agile in body and intelligence (than industrial robots), and in some cases easier for nonprofessionals to program. Analogous to humans, some cobots learn by imitating humans or through guided experience.[7]

---

[4]Universal Robots (2019). Beyond the Cobot Buzz: A Cheat Sheet on How to Choose Between Collaborative and Traditional Industrial Robots. https://info.universal-robots.com/cobots-vs-traditional-industrial-robots [accessed on April 2, 2020].

[5]Cory Roehl (2017). Know Your Machine: Industrial Robots vs. Cobots. Universal Robots, https://blog.universal-robots.com/know-your-machine-industrial-robots-vs.-cobots [accessed on April 2, 2020].

[6]Fortune Business Insights (September 2019). Machinery & Equipment / Industrial Robots market. www.fortunebusinessinsights.com/press-release/industrial-robots-market-9257 [accessed on April 2, 2020].

[7]Smith, K (March 29, 2017). "Cobot" with Deep Learning and Gesture Recognition Hits Audi Production Floor. https://www.allaboutcircuits.com/news/collaborative-robot-deep-learning-gesture-recognition-audi-brussels-factory/ [accessed on June 14, 2020.]; see also Wang, W., Chen, Y., Li, R., & Jia, Y. (2019). Learning and Comfort in Human–Robot Interaction: A Review. *Applied Sciences*, *9*(23), 5152.

Traditional industrial robots require large capitalization costs; individual cobots are relatively inexpensive and thus can scale with the size of the business needs. Cobots are thus transforming many industries, sectors, and regions. Although the market for collaborative robots is currently smaller than the market for industrial robots, the expected growth rate is much higher: The global market size for collaborative robots was $1.57 billion (USD) and is expected to grow to $23.59 billion by 2026.[8] *Universal Robots* currently dominates the collaborative robot market with approximately 60% market share, but companies such as *ABB*, *Robert Bosch*, *KUKA*, and *FANUC* are now competing for this important and consequential market.[9]

Cobots will transform work in large and small organizations. Cobots are currently involved in laparoscopic surgery, providing massages, and disinfecting udders on cows before and after milking.[10] GROWBOT (Grower-Reprogrammable Robot for Ornamental Plant Production Tasks) is a research project at King's College that is using imitation learning to teach cobots to perform small, delicate horticultural tasks such as "taking and inserting cuttings, grading and collating plant specimens."[11]

To illustrate some of the important challenges and benefits of cobots, the following three subsections explore how cobots are used in three different contexts: search-and-rescue, surgery, and order-fulfillment.

---

[8]Fortune Business Insights (November 2019). Machinery & Equipment / Collaborative Robots market. https://www.fortunebusinessinsights.com/press-release/collaborative-robots-market-9395 [accessed on June 14, 2020].

[9]EY-Mint Emerging Technologies Report 2019. Emerging Technologies: Changing how we live, work and play. www.slideshare.net/eraser/emerging-technologies-changing-how-we-live-work-andplay-eymint-emerging-technologies-report2019 [accessed on April 2, 2020]. The statistic regarding market share was cited in this report as originating from CB Insights.

[10]www.ien.com/automation/article/20849060/collaborative-robots-are-showing-up-in-the-strangest-places [accessed on April 2, 2020].

[11]https://horticulture.ahdb.org.uk/project/growbot-grower-reprogrammable-robot-ornamental-plant-production-tasks-phd-studentship [accessed on April 2, 2020].

Search-and-Rescue

In their comprehensive review and analysis of *search-and-rescue robots*, Williams et al.[12] divide search-and-rescue into four essential tasks: search, extraction, evacuation, and treatment. Different robots have been designed and deployed for each of these with little or no overlap between these tasks (although some may be designed to work in tandem with other robots).

Search robots do not need to be social; among the four categories of search-and-rescue robots, search robots (especially airborne and submersible) are the most widely used. They need to be able to survey affected regions and to search for humans and identify unstable environments (such as toxic or volatile gasses and insecure infrastructure), and they need to be lightweight, energy-efficient (enabling nearly continuous operations), and modular (able to easily add or remove diverse types of sensors and robotic arms and manipulators). The basic design imperative is that every second matters—the sooner someone is rescued, the more likely their survival. Discovering humans may require integrating data from various sensors such as CO_2 detectors, heat imaging and visible spectrum cameras,[13] and chemical (olfactory) detectors. In many ways, search robots imitate and expand on what trained search-and-rescue dogs have done for years.[14]

[12]Williams, A., Sebastian, B., & Ben-Tzvi, P. (2019). Review and Analysis of Search, Extraction, Evacuation, and Medical Field Treatment Robots. *Journal of Intelligent & Robotic Systems*, 1-18.

[13]Farooq, N., Ilyas, U., Adeel, M., & Jabbar, S. (2018, October). Ground Robot for Alive Human Detection in Rescue Operations. In *2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)* (Vol. 3, pp. 116-123). IEEE.

[14]Mochalski, P., Ruzsanyi, V., Wiesenhofer, H., & Mayhew, C. A. (2018). Instrumental sensing of trace volatiles—a new promising tool for detecting the presence of entrapped or hidden people. *Journal of breath research*, *12*(2), 027107.

Many search robots are remote controlled, but ideally, they should be able to continue operations and quickly react to local conditions even when communications are disrupted or too slow for real-time responses. Search robots may also need to be sufficiently agile and small enough to enter and navigate small, unevenly shaped openings and tunnels. According to Williams et al., search robots can be divided into three major categories: evaluating the damage and stability of structures and infrastructure, acquiring data for additional processing, and discovering trapped or injured persons. In addition, some of these robots can make minor repairs or adjustments (such as stopping gas flow by turning a valve) or can transport small amounts of medicine and other supplies.

The other three categories of search-and-rescue support extraction, evacuation, and medical treatment. In these activities, interaction with human patients is critical and difficult. Some extraction systems work only if the injured person can be lifted onto a flexible stretcher that is attached to the robot (e.g., *iRobot Valkyrie*[15]); in other cases, the robot can lift the victim (e.g., the *Battlefield Extraction-Assist Robot*, which is produced by Vecna Robotics[16]), but this is advisable only if the victim has had no head or neck trauma.

Evacuation systems, such as *LSTAT* (Life Support for Trauma and Transport), transport patients to field hospitals for further medical support and typically monitor the patient's blood pressure, pulse, temperature, oxygen levels, and so on. *LSTAT* has been successful in hospital and military field studies.[17]

---

[15]Valkyrie was developed by iRobot in 2003 as a recovery robot. See `http://robotfrontier.com/gallery.html` [accessed on April 2, 2020].

[16]BEAR is a humanoid robot developed by Vecna Technologies in 2004.

[17]Williams, A., Sebastian, B., & Ben-Tzvi, P. (2019). Review and Analysis of Search, Extraction, Evacuation, and Medical Field Treatment Robots. *Journal of Intelligent & Robotic Systems*, 1-18.

Surgery

This subsection explores the use of cobots during surgery. In current surgical applications, the cobot is remote controlled by a human surgeon.

The use of remote-controlled robots in medical surgery is not without controversy. The surgeon monitors multiple displays and instructs the robot. The robot, which is often situated above the operating table, uses precision, sensory-laden "fingers" to examine and manipulate.

When designing a cobotic system, it is important to ask psychological and sociological questions, in addition to the standard questions about effectiveness, functionality, reliability, and cost. Is the resulting human-machine system an effective and healthy partnership? What happens when remote-controlled robots are used in surgery? How does power and authority shift? Does it alter the patient-doctor relationship? Does it modify the relationship between the surgeon and other surgical team members?

According to one study (Juo et al., 2018),[18] which examined the doubling of robotic-assisted laparoscopic surgeries between 2008 and 2013: "No significant association existed between the frequency of robotic-assistance usage and relative outcome statistics such as mortality, charge, or length of stay." No difference in self-reported postoperative outcomes was also reported for robotic and traditional laparoscopic surgery for women with endometrial surgery. Both laparoscopic techniques led to better outcomes than open surgery, but no differences in patient-reported outcomes were found between the robotic and traditional laparoscopy.[19]

[18]Juo, Y. Y., Mantha, A., Abiri, A., Lin, A., & Dutson, E. (2018). Diffusion of robotic-assisted laparoscopic technology across specialties: a national study from 2008 to 2013. *Surgical endoscopy*, *32*(3), 1405-1413.

[19]Ferguson, S. E., Panzarella, T., Lau, S., Gien, L. T., Samouëlian, V., Giede, C., ... & Bernardini, M. Q. (2018). Prospective cohort study comparing quality of life and sexual health outcomes between women undergoing robotic, laparoscopic and open surgery for endometrial cancer. *Gynecologic oncology*, *149*(3), 476-483.

However, in a recent comparison of conventional laparoscopic and robotic-assisted bariatric surgery, robotic surgery was associated with significantly longer operations, and "in gastric bypass, rates of aggregate leak and bleeding were higher with robotic surgery, while transfusion was higher with laparoscopy." In sleeve gastrectomy cases, other outcomes, such as "… reoperation, readmission, sepsis, …," were higher with robotic surgery.[20]

A recent meta-analysis of 27 clinical reports of robotic and traditional laparoscopies,[21] dated from 1981 to 2016, found that outcomes in robotic-assisted methods were not significantly better than traditional methods, apart from lower *estimated* blood loss. Indeed, traditional methods resulted in better operative times and reduced complication rates and overall cost.

Robotic surgery is a transformative technology that reorganizes how teamwork is achieved. The use of these robots has radically altered the surgical environment. In human-conducted surgery, the surgeon works directly on the patient, peering directly into patient. The surgical team hovers around the surgeon, orienting themselves to his or her preferences. Traditional laparoscopy changes this somewhat, but the surgeon and the surgical team are still close to the patient and the surgical team is still oriented around the surgeon.[22]

---

[20] Acevedo, E., Mazzei, M., Zhao, H., Lu, X., & Edwards, M. A. (2019). Outcomes in conventional laparoscopic versus robotic-assisted revisional bariatric surgery: a retrospective, case-controlled study of the MBSAQIP database. *Surgical endoscopy*, 1-12.

[21] Roh, H. F., Nam, S. H., & Kim, J. M. (2018). Robot-assisted laparoscopic surgery versus conventional laparoscopic surgery in randomized controlled trials: a systematic review and meta-analysis. *PloS One*, *13*(1), e0191628.

[22] Pelikan, H. R., Cheatle, A., Jung, M. F., & Jackson, S. J. (2018). Operating at a Distance - How a Teleoperated Surgical Robot Reconfigures Teamwork in the Operating Room. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 138.

Robotic surgery changes this: A many-armed surgical robot, such as the *da Vinci* telesurgery robot, is positioned over the patient. The robot's arms and tiny articulators easily move with near-perfect precision and stability where humans cannot. The sensors attached to the articulators provide information about the patient's condition that is not possible in traditional surgery. And the AI that ingests the sensory information provides real-time predictions about the health of the patient, such as the likelihood that the tissue under observation is cancerous or benign. These are great achievements of robotic and machine-learning technologies.

However, the surgeon is now located father from the patient. In another part of the room, the surgeon huddles over a monitor and control console, controlling the robot arms through various gestures and commands. The surgical team is no longer pressed close to the surgeon, but some of them are monitoring the patient. Despite the earlier mentioned benefits, robotic surgery transforms the physical arrangement of the surgical theater and the way in which humans interact. The need for verbal communication is increased because the humans can no longer communicate easily by glance and gesture[23] and students don't participate as extensions of the surgeon's physical and sensory system and therefore don't get the same level of direct training.

Evaluations of robotic surgery should not only measure traditional key performance indicators like operative duration, cost, blood loss and other complications, and postoperative health, but also the sociological impact: how does the inclusion affect the performance of the surgeon and the operative team? The training of students? And the real-time decision-making capabilities of the team? The longer operative times observed in many robotic-assisted surgeries may be due to the intra-team

[23]Pelikan, H. R., Cheatle, A., Jung, M. F., & Jackson, S. J. (2018). Operating at a Distance - How a Teleoperated Surgical Robot Reconfigures Teamwork in the Operating Room. *Proceedings of the ACM on Human-Computer Interaction, 2*(CSCW), 138.

communication obstacles introduced by the placement and actions of the robot. This might improve with time and experience, as the surgical team adapts to their new "member," but some of the effects may have long-term consequences for the training and behavior of the surgical staff.

Warehouses

Current robotic-assisted order-fulfillment warehouses offer other insights into the benefits and pitfalls of collaborative robots. In December 2018, as humans and robots were quickly processing customer orders in a New Jersey Amazon warehouse, a can of bear repellent was accidentally punctured by a robot.[24] Fifty-four employees were exposed to the noxious fumes, of which two dozen were hospitalized. A similar accident occurred in 2015, at an Amazon facility in Haslet, Texas, when a robot rolled over a can of bear repellent.[25]

These accidents raise questions about the use of robots; human workers have also dropped bear repellent, but a human stepping on a can of bear repellent will not cause an explosion; instead, it will typically cause the human to stop pressing their foot, to bend down, and to restock the undamaged can. Developing an efficient, cost-effective solutions that allow robots to learn to monitor the consequences of their actions, to halt actions that might have dire consequences, and to take corrective action is nontrivial and a research challenge.

Retrieving fallen items is just one reason why a human worker might enter a restricted space. Other reasons for entering an area in which

[24] Louise Matsakis (December 6, 2018). This Wasn't Even Amazon's First Bear Repellent Accident. Wired, www.wired.com/story/amazon-first-bear-repellent-accident/ [accessed on April 2, 2020].

[25] Brian Heater (January 18, 2019). Amazon built an electronic vest to improve worker/robot interactions. TechCrunch. https://techcrunch.com/2019/01/18/amazon-built-an-electronic-vest-to-improve-worker-robot-interactions/ [accessed on April 2, 2020].

robots are working include setup, maintenance, and testing. The danger for humans working in these restricted spaces is obvious, and Amazon recently introduced an electronic vest to improve human and robot interactions. The vest allows robots to detect human presence more quickly and accurately and to move along paths that avoid humans. Despite the risks, the use of robots in warehouses improves efficiency and accuracy—robots can excel in tasks that require superhuman strength and endurance.

Working Alongside Humans

Each of the previous examples (search-and-rescue, surgery, and factory teams) reflects the benefits and dangers that can occur when humans and robots work in proximity. Cobots can be larger, stronger, and more agile. They can be privy to data that humans cannot access in real-time and can respond to that data much more quickly than humans. Humans on the other hand can be more adaptive to changes in the environment and more flexible in how they think about and resolve problems. Robots tend to be constructed around specific, well-defined objectives; humans evolved to be adaptive. So how can they cooperate in human-robot teams that benefit human society?

Humans are social animals and the way in which we frame our interactions with our tools, other species, and our environment is strongly influenced by the way we interact with one another. We anthropomorphize. Some give names to their cars and most give names to their pets. Humans model their interactions with computers and other intelligent technologies on human-to-human dialogue, which is a good strategy because the interactions methods used by these technologies were designed by humans, and can be viewed as extended, computer-mediated, dialogues. As robots become more intelligent and more varied in their responses, we will increasingly ascribe personalities and names and anthropomorphize their behavior.[26] We will adapt to their limitations and abilities, and through iterative design and machine learning, they will adapt to us.

Levels of Automation and Collaboration

Automation and collaboration are not simple concepts that are either absent or present in a human-machine interaction. It is a continuum that exists across multiple dimensions. In Table 4-1, we present our method describing this continuum,[27] by decomposing it into automation vs. human-machine collaboration (the second and third columns) and into five levels of collaboration (0 to 4) from interacting with non-intelligent mechanical tools to collaborating with robots that have artificial intelligence and autonomy. The column labeled, *Automation*, provides examples of machines that operate without frequent human supervision. The *Human-Machine Collaboration* column focuses on the examples with frequent communication between humans and machine—the behavior of each, constraining the other.

[26]Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press.

[27]This table is influenced by prior research on levels of automation and autonomy, particularly, Abbass, H. A. (2019). Social integration of artificial intelligence: functions, automation allocation logic and human-autonomy trust. *Cognitive Computation*, *11*(2), 159-171; Endsley, M. R. (2017). From here to autonomy: lessons learned from human–automation research. *Human factors*, *59*(1), 5-27; and the National Highway Traffic Safety Admin (NHTSA). (2013). US Department of Transportation, preliminary statement of policy concerning automated vehicles. *NHTSA preliminary statement*. www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf [accessed on April 2, 2020].

***Table 4-1.** Levels of Automation and Collaboration*

Level	Automation	Human-Machine Collaboration
0. *Non-intelligent tools*	*No intelligent automation* • Non-intelligent devices that operate with little or no human supervision. • Examples: gas engines, boilers, water turbines that existed prior to 1930s.	*No intelligent interaction* • Human makes all decisions, and interpretation is fixed. Examples: driving or braking in cars made before the 1970s; mechanical looms.
1. *Human-directed interactive tools*	*Human-directed automation* • Fixed logical process initiated and designed by humans. • Examples: batch processing and RPA; industrial robots working in restricted areas; the Jacquard loom with punch cards (c.1801).	*Human-directed interaction* • Human makes all decisions; machine can make local adjustments. Examples: modern antilock braking systems and cruise control in cars; standard text (w/ autocorrect) or graphic editors.

(*continued*)

Table 4-1. (*continued*)

Level	Automation	Human-Machine Collaboration
2. ***Partial or conditional collaboration***	*Human-assisted automation* • Human selects goals, robot recommends actions and once confirmed acts with limited autonomy. • Examples: smart buildings automatically adjusting lighting and airflow; intelligent process automation (IPA).	*Machine-assisted interaction* • Human selects goals, receives continuous feedback, and can quickly assume full control. Robot has limited autonomy. • Examples: virtual assistants that reserve flights for air travel; traffic-aware cruise control; remote-controlled surgical robots.
3. ***High collaboration***	*High automation* • Humans (who can be remote) determine goals and monitor situation. Robot acts with autonomy. • Examples: autonomous vehicles with relaxed human supervisor (not possible today except in experiments).	*Human-directed cobot teams* • Humans determine goals, and human and robot coordinate actions. • Examples: search-and-rescue cobots; robots that work alongside humans in warehouses.
4. ***Machine-directed governance & coordination***	*Full automation* • Robots determine mission and perform all necessary actions (not possible today).	*Machine-directed cobot teams* Robots determine mission and direct human-robot teams (not possible today).

We do not represent in Table 4-1 the so-called singularity in which machines reign supreme and are supremely intelligent and autonomous (and often tyrannical),[28] although *Level 4* may hint at this possibility.

Level 0 describes the early days of human-machine interaction (precomputers). This is the age of mechanical and electromechanical machines, in which many humans designed, created, operated, and maintained these machines. Some of these machines were complex and required constant human attention for their operation (such as early automobile), and the relationship between input and output was analog and continuous—turning the steering wheel slightly, turned the car slightly). Collaboration, at this level, is between humans; machines are simply tools or infrastructure that support human interactions and physical needs. Technology has continued to evolve Level 0 machines (such as traditional refrigerators and furnaces) that can operate without human supervision for long periods of time.

Level 1 reflects a monumental shift in how humans and machines interact. Mainframe computers, which emerged in the decade following World War II, receive and act upon human instructions in batch mode (automation).[29] The Jacquard loom with its punch cards for input (c. 1801) was an early form of Level 1 automation, and robotic process automation (RPA) is currently a popular form of Level 1 automation.

Human-directed tools became more interactive (human-machine interaction) and personal during the 1960s and 1970s. During the period cruise control and antilocking brakes became popular in cars. Personal computing and the highly interactive software (human-machine interaction) first emerged as a mass-market consumer device in the 1970s. Although not truly collaborative, these machines required designs

[28]Cadwalladr, Carole. "Are the robots about to rise? Google's new director of engineering thinks so." *The Guardian* 22 (2014).

[29]Batch mode computing can operate with minimal interaction and execution time and resources can be scheduled.

that anticipate human behavior and made minor adjustments based on human input and environmental conditions. Standard text editors with autocorrect are a good example of Level 1 collaboration.

Level 2 represents another major advance in human-machine interaction, involving (a) inexpensive and efficient data storage and processing, (b) access to global data, (c) miniature Internet of Things devices, and (d) new machine-learning techniques—deep learning in particular. This has enabled powerful techniques for partial or conditional machine intelligence in:

- *Automation: A*lgorithms and systems for intelligent process automation (IPA). IPA is a predefined combination of "business rules, experience-based context determination logic, and decision criteria to initiate and execute multiple interrelated human and automated processes in a dynamic context."[30] Like RPA, but with much greater intelligence and conditional logic, IPA delivers complex series with little or no human assistance.

- *Human-machine collaboration*: Algorithms for gesture and natural language recognition, and chatbots and personal assistants that explore complex databases, such as air flight *information*. At the device level, IoT microprocesses, sensors, and tactile feedback interfaces have supported traffic-aware cruise control and remote-controlled surgical robots.

For the present discussion, chatbots and personal assistants are particularly interesting because they constitute the emergence of intelligent one-to-one, or dyadic interaction between human and

[30]2755-2017—*IEEE Guide for Terms and Concepts in Intelligent Process Automation.* https://standards.ieee.org/standard/2755-2017.html [accessed on April 2, 2020].

machine. We can have short directed conversations with personal interactive software like Amazon Alexa, Apple Siri, Google Assistant, and Microsoft Cortana. They have been designed through rules and/or machine learning to emulate humanlike responses. Collaboration is dyadic—two actors, one human and one machine. Notably, their behavior is not autonomous—they act within very narrow, predefined limits. A human personal assistant, who can think "out-of-the-box," might help you with your travel plans by suggesting a city that you had not considered or might even suggest a *staycation*, in which you vacation at home. But, as of 2020, we are not aware of any commercially available chatbot that interrupts a flight reservation dialogue to suggest that the caller should consider a vacation in a different city or simply stay home.

Level 3 recognizes a critical milestone in robot intelligence and human-machine interaction—autonomous robots. At this level, humans and machine work together monitoring each other's behaviors and acting to minimize risk and maximize defined benefits.

Automation at this level is termed *cognitive automation* which is defined as a system that achieves its objectives by performing "corrective actions driven by knowledge of the underlying analytics tool itself, [and iterating] its own automation approaches and algorithms."[31] It can rewrite itself! For some, this a *tipping* point in human-machine interaction and inevitably leads to *Level 4*, in which machines dictate our behavior and objectives (this might be limited to specific situations such as search-and-rescue or mining operations).

Human-machine collaboration at *Level 3* fully incorporates the use of cobots in work teams. These cobots can observe and interact with multiple actors (humans and machines) and fuse this data into a coherent model of their social environment. To achieve this level of integration, there are physical and emotional cognitive challenges for both humans and robots.

[31]2755-2017—IEEE Guide for Terms and Concepts in Intelligent Process Automation.

Clearly, robotic hardware needs to be designed for high interactivity with humans and must minimize the likelihood of harm to humans. But there are cognitive and emotional obstacles. What does it take to be a good team member? We shall explore this question in the next section.

We shall not consider *Level 4* in more detail in this chapter, reserving that for discussion in the concluding chapter of this book.

Teamwork: From Conversational Interfaces to Physical Cobots

The remainder of the chapter will focus on *Level 3*, *High Collaboration*, and the evolution from conversational interfaces that interact with one person at a time to cobots that collaborate in teams of other cobots and humans.

The challenge of creating chatbots and virtual assistants that are truly aware of multiple human teammates has been examined by Seering et al.[32] In their systematic classification of research on chatbots and on deployed chatbots, Seering et al. concluded that chatbots that can participate in a multiperson is an important but underresearched topic: "None of the chatbots described in the research literature were designed to be members of a community, but rather they were all designed as tools to support their communities."[33]

---

[32]Seering, J., Luria, M., Kaufman, G., & Hammer, J. (2019, April). Beyond Dyadic Interactions: Considering Chatbots as Community Members. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 450). ACM.

[33]In their systematic survey of research on chatbots, Seering et al. identified 104 research papers, of which 91 were primarily concerned with dyadic communication, 6 dealt with broadcasting chatbots (chatbots that send one-way messages to many recipients), 6 focused on multiuser chatbots, and 1 paper did specifically concentrate on any of these categories. Among the 130 chatbots that they identified that are used outside of academic research, 103 were dyadic, 14 were broadcasting chatbots, and 13 were multiuser. Of the 13 multiuser chatbots, 11 were used on online community platforms to host chatroom-style interactions (e.g., Twitch, Discord, and Slack).

Although there are many valuable applications that require only interacting with, and on behalf of, a single person (e.g., a travel reservation chatbot), physical robots that are intended to move in the real world among other humans should be able to coordinate their physical, social, and communicative behaviors in the context of other robots and multiple humans. Context and familiarity can reduce the inherent complexity.[34]

In 2017–2018, the authors were part of research team whose goals were to develop computational models for task-oriented human-robotic systems. This research was led by Professor Moncef Gabbouj and his students at Tampere University. The authors (who at that time were research scientists at *CA Technologies*) and M. Vakkuri (from *Tieto Oyj*)[35] provided the business framework and industrial constraints for modeling the interactions.

One context (or scenario) developed to guide this research was search-and-rescue. Robots were tasked with helping humans navigate a dangerous environment. This context can easily be extended to business settings in which humans and robots interact in factories, shipping docks, and so on. In addition to their other tasks, the robots would need to classify environments as safe or unsafe for robots and humans and, if safe, to initiate various activities. A schematic of this problem is shown in Figure 4-1. As can be seen in this figure, in order to classify the environment as safe or unsafe, the robot(s) would need to

[34]For example, an autonomous vehicle needs to coordinate its actions with those of other vehicles and might only verbally interact with the driver. A robot that is a part of a team of dock workers might only take commands from certain team members and might only carry out certain tasks. Other humans in these situations will over time learn how these robots interact and will likely adapt to what can be expected from the robots and how to avoid interfering with their proper behavior.

[35]Tieto is now called *TietoEVRY*, following a merger in 2019.

1. Scan the environment and other data streams
2. Identify and track humans, robots, machines, and other aspects of the environment
3. Update a situation model based on (a) the results of steps 1 and 2, (b) earlier situation models, if they exist; (c) other information about humans, robots, and machines; and (d) other information about the environment and its physical properties

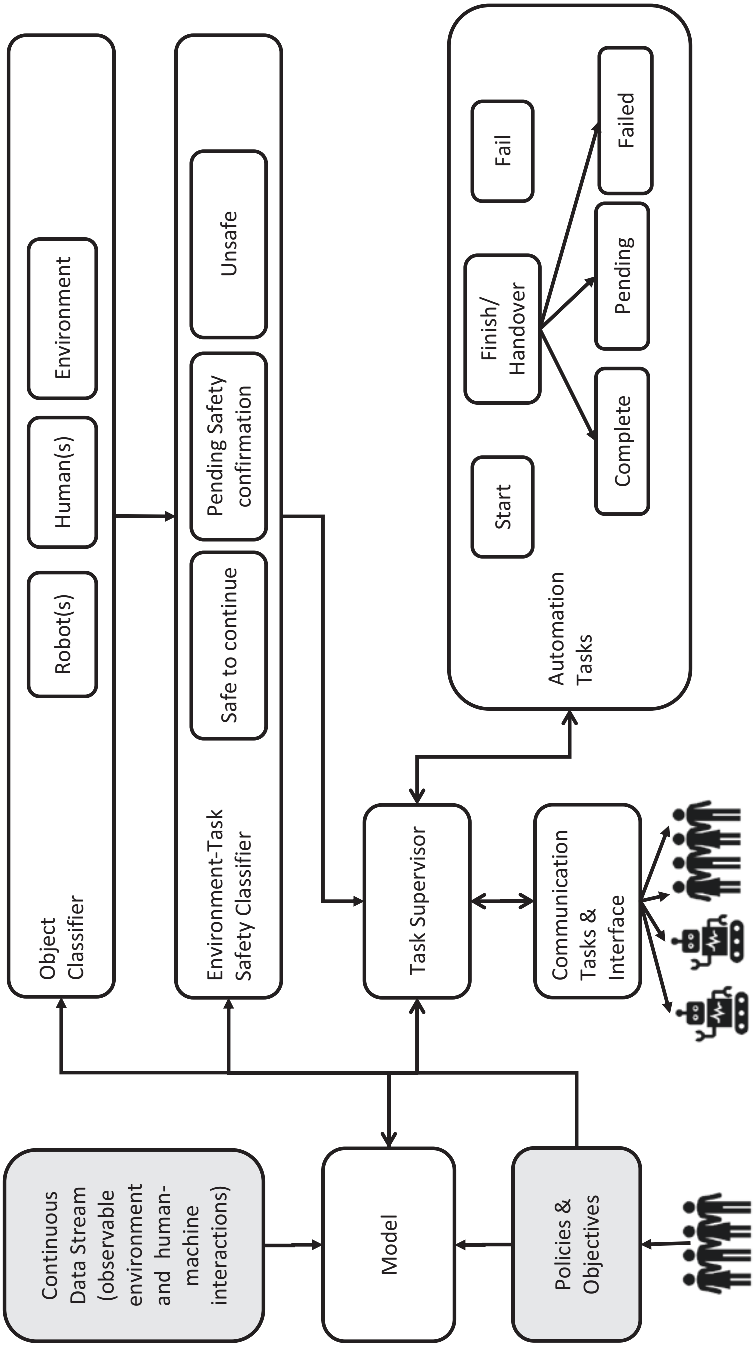

Figure 4-1. *Cobot software architecture. The inner boxes represent data that is external to the architecture. The outer boxes indicate software that is part of the overall cobotic system*

Overall cobot control could be either (a) individual-based in which case each cobot forms a unique model of the environment, (b) distributed in which case the cobots communicate with each other to form a unified model, or (c) centralized in which case a control server forms the model of the environment using inputs from each cobot and then instructs the cobots.

If no humans or robots currently existed in that region, then the classification would be used to decide whether humans or robots (and what type of robots) could enter the space. If humans and robots already inhabit the region, then the classification can lead to a decision to evacuate or continue working. If the decision was to continue working, then the robots could carry out their tasks, using the same situation model that formed the basis for the safety analysis.

To accomplish its safety analysis and to be able to navigate its physical and social environment, cobots functioning at Level 3 (High Collaboration) would require an effective and efficient algorithm for combining disparate sources of data to create a coherent situation model that supports collaborative search-and-rescue. The challenges of combining different sources of data are explored in Chapter 6, "Robots in a World of Data," but its relation to collaboration should be emphasized, here. In order to appropriately interact with other robots and humans, cobots must integrate (or fuse) information about the physical and social environment: the safety and capabilities of the physical infrastructure, the location of the robots, humans and other objects in that environment, and what the other robots and humans are doing and communicating. The sophistication and complexity of this information will depend on the roles that the cobot is expected to fulfill.

Conflicts and Trust

In almost any multihuman situation, conflicts arise. Adding semi-autonomous collaborative robots to these situations is unlikely to reduce the number of conflicts. Some conflicts arise because the most efficient path between the current location of a human, or cobot, and a desired destination is blocked or about to be blocked by another human or cobot (or anything else in motion). There are many studies of this type from air traffic management systems[36] to robots in a shared workspace.[37] Other conflicts arise because two agents wish to use the same object: they are attempting to grab the same box, edit the same document, or make use of another robot or person. In these cases, conflicts arise because different agents are attempting to use the same physical or virtual *extrinsic* resource.

Another type of conflict arises because two or more agents have a different perspective.[38] This might be due to differences in knowledge, belief bias, experiences, commitment, or ways of reasoning or acting. For example, if two robots are to carry a large object and they move at different speeds or heights, there is a conflict that must be resolved. You can observe humans in this type of conflict by watching movers carry furniture up a staircase. These are *intrinsic* conflicts.[39] They arise because of differences between individual agents (or even within an agent).

---

[36]Tomlin, C., Pappas, G. J., & Sastry, S. (1998). Conflict resolution for air traffic management: A study in multiagent hybrid systems. *IEEE Transactions on automatic control, 43*(4), 509-521.

[37]Wong, K. W., & Kress-Gazit, H. (2015, May). Let's talk: Autonomous conflict resolution for robots carrying out individual high-level tasks in a shared workspace. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 339-345). IEEE.

[38]Tessier, C., Chaudron, L., & Müller, H. J. (Eds.). (2006). *Conflicting agents: conflict management in multi-agent systems* (Vol. 1). Springer Science & Business Media.

[39]Castelfranchi, C. (2015). The cognition of conflict: ontology, dynamics, and ideology. In *Conflict and Multimodal Communication* (pp. 3-32). Springer, Cham.

From a design perspective, there are three basic strategies for conflict management: avoidance, detection, and resolution. If the designer of a cobotic system (cobots + humans + environment) can foresee all possible conflicts, then these can be explicitly embodied in a set of rules or implicitly in a deep-learning system through repeated exposure during training to the conflict situations. However, overengineering a situation to reduce conflicts can have negative consequences as well. As Easterbrook states, "not only is conflict inevitable in society, both within and between individuals and organizations, but that conflict has a useful role in facilitating change and producing higher quality group decisions."[40]

In lieu of completely predictable interactions, cobots and any supervisory system need mechanisms for identifying and resolving the conflict. There are many strategies for resolving conflict and most of them involve understanding why the conflict arises and how the goals and perspectives differ. Conflict resolution requires trust and an evaluation of common ground—the common knowledge and beliefs shared by all participants.

Klein et al.[41] argue that collaboration, or joint activity, requires that each agent must agree to a mutual intention to work together. They must also be predictable and responsive to each other and must work to maintain common ground. These allow teams "to facilitate coordination, work toward shared goals, and prevent breakdowns in team coordination," and they are fundamental to trustworthiness. Trust evolves over time, but humans quickly form and reform judgments about trustworthiness, based on their short-term interactions with others.[42]

---

[40]Easterbrook, S. (1991). Handling conflict between domain descriptions with computer-supported negotiation. *Knowledge acquisition*, *3*(3), 255-289.

[41]Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems*, *19*(6), 91-95.

[42]Greenspan, S., Goldberg, D., Weimer, D., & Basso, A. (2000, December). Interpersonal trust and common ground in electronically mediated communication. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 251-260). ACM.

For cobots to be treated as a team member, the policies that govern their behavior and accumulated data must be transparent and quickly comprehensible to the other team members. Similarly, while they are working in a team, cobots must ensure that others can easily predict and, if necessary, adjust the cobot's behavior.

In successful human teams, members value and protect each other's privacy and look out for one another. *If a cobot is present on a team, and a team member says or does something that violates business policy, should the robot issue a report to management?* This is challenging ethical question, and creating a policy for cobot behavior is nontrivial. A comparison to controversies surrounding email highlights some of the important concerns. All email on a corporate server is legally owned by the corporation. Corporations could proactively analyze emails for indications of unhappy (and potentially harmful) employees, for evidence of romantic partnerships that violate company policy, and for suggestions of ethical violations. However, most companies refrain from this, respecting the privacy of an individual's email.

Humans that work with cobots will need to be assured that the cobot is not recording every action and utterance or, if so, that the data will be kept private unless there is an extraordinary and compelling legal reason to analyze and expose it. It might be useful to have detailed recordings of every action taken in a surgical operating room, but having the cobot do so might affect how the human medical staff behaves. Cobotic systems are not strictly hardware and software implementations. The roles and objectives of cobots must be aligned with social expectations for ethical interactions.

Guidelines for Designing a Cobot

Cobots can be conceived of as decision-making entities that operate through prediction and automation but are constrained by interactions with the physical and social environment. Apart from the physical requirements which are beyond the scope of this book, a cobot that works on a team with humans should be able to[43]:

1. *Integrate data* that it receives from scanning the environment (across multiple sensory modalities) in order to:

    a. Identify the active participants and other movable objects in the environment.

    b. Perceive the status and intentions of the participants. This includes being able to identify requests for help, follow directions to attend to something or someone in the environment, and obey human-provided instructions to modify its behaviors or objectives.

    c. Perceive unexpected patterns of data, recognize that they are unexpected, and react accordingly.

2. *Construct models* of their environment appropriate to the team's goals. The cobot must be able to construct, maintain, and modify its models of the current physical and social environment. To negotiate a social and physical space, robots will need to classify the multilayer roles that humans

[43]This list is influenced by the work of Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., and Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems*, *19*(6), 91-95.

assume in specific contexts. This is associated with the conversational roles that humans adopt but includes *the* challenge of categorizing human activities to ascertain a human's role in a task (e.g., decision-making roles in a team may switch back and forth among human experts as an evolving situation demands different skills).

3. *Apply policies* to regulate its behaviors and resolve conflicts. Notably, as cobots become more autonomous, they also become less predictable. Policies restore some of that predictability by setting limits and governing actions. In line with these policies, the cobot must be able to dynamically detect possible failures and create plans to avoid them while maintaining the overall team goals.

4. *Complete tasks* by integrating (or fusing) models of physical and social environment with policies and objectives (as defined by designers, integrators, and supervisors). This includes being able to negotiate goals and be (re)directed by humans and other cobots through:

   a. *Physical and conversational turn-taking and activity management*: The cobot must determine when it is its turn to contribute an action, or when it can interrupt an ongoing conversation.

   b. *Attention management and joint attention*: The cobot must be able to determine what others are attending and must be able to direct the attention of others when necessary. For example, an autonomous vehicle should be able to direct the driver's attention to road obstacles that require the driver's decision, and a search-and-rescue cobot should be able to direct the team's attention to survivors or teammates that require help.

c. *Goal-directed actions*: Cobots must be able to initiate actions either directly in a physical environment (through manipulation and movement), or indirectly through data and human forms of communication (gestures, text, speech). This includes interactions with other robots and humans that promote both task objectives and team-building goals. For example, in a factory setting, some of these actions may be directed to transporting product components from one section of the factory to another, and some actions may be social greetings to humans that the robot encounters along its path. The social greetings help to create trust that the robot is functioning, awareness that the robot is present and carrying out a specific task, and perhaps also relay specific messages.

d. *Task completion*: Cobots should complete their tasks within expected time and resource constraints or indicate to the appropriate stakeholders when delays or failures might occur.

As we have observed, the design of successful cobotic systems (cobots, environment, humans, policies, and objectives) is more than just the physical and software construction. Analyzing and optimizing a cobotic system is much like analyzing and optimizing a work environment in which multiple intelligent beings interact. Careful attention must be given to team dynamics and activities[44]:

[44]These questions are derived from activity theory. See Engestrom, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, 43(7), 960-974.

- What are the objects of value that are created by the team (documents, practices, transactions, physical objects, etc.)?
- Who are the stakeholders (not just the team) in this value creation? Who has power? Who is considered part of the team, and who is not?
- How are the activities distributed across the participants and what happens when there are conflicts?
- What tools and practices are required for team's activities and objectives, and who controls them?
- What are the rules that guide objective setting, task management, and conflict resolution? And how is trust maintained or repaired?

In their analysis of incident response teams in computer and network operations centers, Brown, Greenspan, and Biddle[45] found that team roles and composition were fluid—different members would come get involved at different times and might belong to multiple teams. Moreover, their role definition might change as the teams evolve. As cobots evolve to become assistants in complex operations, they will also need to evolve organizational fluidity.

[45]Brown, J. M., Greenspan, S., & Biddle, R. (2016). Incident response teams in IT operations centers: the T-TOCs model of team functionality. *Cognition, Technology*.

Summary and Conclusion

In the beginning of this chapter, we examined the benefits that collaborative robots are bringing, or might bring, to search-and-rescue, surgical, and warehouse operations. We also considered some of the technical, organizational, and ethical challenges that could weaken these benefits. For example, the social and physical dynamics between the surgeon and the surgical team are transformed and perhaps weakened by using large remote-controlled surgical robots that are positioned over the patient.

We then considered what it means to collaborate with others and introduced a five-level model of collaboration:

1. Non-intelligent tools
2. Human-directed interactive tools
3. Partial or conditional collaboration
4. High collaboration
5. Machine-directed governance and coordination

Many of today's robots and automation software are examples of either *human-directed interactive tools* or *partial or conditional collaboration*. We are on the verge of high collaboration between man and robots. Even today, semi-autonomous cars and jets make moment-to-moment decisions without human input, and robotic ships that map the ocean floor devise strategies and tactics for carrying out their missions. To achieve greater levels of collaboration, we must construct robots that are able to participate in teams, as a team member. These cobots will not be fully human, but they will be more than a simple tool.

The discussion, therefore, turned to question of what's required for team collaboration; what social and cognitive abilities are required? And how can a machine resolve the conflicts that typically arise in teamwork?

To answer these questions, we summarized research on social collaboration and teamwork and provided guidelines for designing cobotic systems, including the ethnographic questions that should be discussed with the humans that will interact with cobots.

There is a growing scientific and nonscientific literature about the dangers of intelligent robots, suggesting that robots will take our jobs, become our master, or turn the workplace into places in which humans will need to act more like robots in order to coexist with robots.

The differences between humans and cobots (as they evolve) are partly structural (biological vs. engineered) and partly social construction. As cobots become integrated into our workforce, our expectations, biases, and aspirations will determine our human-machine relationship as much as the physical differences between human and machine. What we design, develop, and use should be based on a clear sense of objectives and societal deliberation, and not just the marketplace pressures for efficiency and low cost.

CHAPTER 5

Robots Without Arms

Smart Buildings and Transport

We are looking toward a future where more things become smart with the addition of many sensors and actuators connected to the world of the Internet of Things (IoT). Software robots and physical robots are being developed and increasingly deployed in the home and workplace with lights, door locks, air conditioning, and cookers increasingly becoming sources of data and action. Software robots are becoming smarter, managing process automation using robotic process automation (RPA) or using chatbots capable of holding a two-way conversation as we saw in Chapter 4, "Robotic Process Automation." Physical robots are working in supply chains, in warehouses, and in the delivery of goods. These have all been discussed in earlier chapters. In this chapter we will examine the current and future impacts on society and two special cases of automation: smart buildings and autonomous vehicles sometimes called driverless vehicles. This chapter will focus on these two cases that could best be described as "robots without arms."

Can smart buildings and autonomous vehicles be called robots? They provide automated responses to the needs of users, but they don't gather or manipulate items in an environment. Robot vacuum cleaners move around their environment and gather items from the floor. Warehouse robots use a variety of rollers, arms, and other manipulators to move items into and out of storage. Smart buildings are stationary with users entering

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_5

or leaving the building. Autonomous vehicles move passengers or goods and the passengers or goods are not gathered by the vehicle but enter under their own steam or by the agency of a third party. The term robots is sometimes used to describe technology that is an integral part of society that supports the activities of humans without being controlled by the users. Robots without arms is a term that can be used to differentiate smart buildings and autonomous vehicles as robotic special cases with their own unique challenges and values to both work and society.

Transport vehicles are ubiquitous and are a general requirement in developed societies. Autonomous vehicles can be classed as a special case because they carry goods and passengers, are bristling with sensors, and fulfill a basic need in society. Goods are delivered by trucks, people are transported by cars or public transport, but there are legal and social limitations on the use of transport. Currently a user has to walk to a bus stop to be driven somewhere. Others who are driving a car have to be over a certain age and capable of driving, measured by a driving test. The elderly may also lose their capability to drive or the means to own a vehicle. In a discussion with an elderly relative, one of the authors suggested they gave up driving. The elderly relative became emotional since they equated the ability to drive and own a car as vital to their independence and well-being. Autonomous vehicles that can transport goods and users from place to place without human intervention are expected to have a positive effect on both business and personal users. The fate of delivery drivers and truckers will be discussed later in this chapter.

The Internet of Things is helping to increase the intelligence of commercial and public buildings to the stage where the building's capabilities such as lighting and air quality are integrated into a single view of the building and are capable of customizing the environment for individuals. This customization can make buildings change the lighting levels for visually impaired workers or increase the humidity in an area used by a worker with eczema. Only a small but increasing number of buildings are capable of this level of customization and integration and they are generally purpose built. Are smart buildings robots? They have autonomy within parameters to change the internal environment, but they don't move. They are human centric with a goal of making life better for the users. Smart buildings are being developed and deployed to enhance the working life of users both in commercial buildings and public buildings. Home use of smart technology relies on the needs and requirements of the owners but is currently patchily implemented. Personal choice influences the use of smart home technology and it is unlikely to require the same scale and direction of smart technology integration and will not be part of this discussion.

Smart Buildings

Many organizations have been looking to make their buildings better places to work in and more efficient in the use of heat and light. This is leading to optimism in the buildings industry with growth figures of 30% in the smart building market or growth from $8.5 billion in 2016 to approximately $58 billion globally in 2022.[1] It is not clear whether these figures include the growth of retrofitting existing buildings or only include new buildings. What is clear is the amount of hype both in the commercial and home market. Corporate brochures are making claims for their company buildings. On television there are advertisements for doorbell cameras, home control apps, and heating controls. One feature common to both markets is the potential for integration of their technologies to enable the use of management portals and consoles to achieve a complete view of all systems that may be interlinked. There are a number of characteristics that make smart buildings a good investment.

[1]Blue Future: The Future of Smart buildings, https://medium.com/@BlueFuture/the-future-of-smart-buildings-top-industry-trends-7ae1afdcce78 [accessed on April 13, 2020].

It is interesting to note that in the Gartner Hype Cycle for Emerging Technologies 2018,[2] smart workspaces were almost at the top of the "peak of inflated expectations." In the 2019 hype cycle,[3] smart workplaces are nowhere to be seen. An explanation for this discrepancy appears at the end of the 2019 emerging trends descriptions. This appears to be a Gartner decision to refocus on emerging trends that have not appeared in earlier versions of the hype cycle and removing trends that are still important but have been featured for a number of years. Some of these trends are not so much a trend but static taking up real estate on the hype cycle and prevent newer more dynamic trends to be noted. It may also be the case that some of these trends have vanished because the world has moved on and left them behind.

Benefits of Smart Buildings

There are a number of benefits of smart buildings that we will discuss here. Benefits include improved building efficiency, lighting improvements, air quality, and temperature and humidity. The advantages of using smart buildings are their potential to improve working life for employees in offices or visitors to public buildings, reducing energy consumption and improving building efficiency, increased productivity, and better use of resources. New buildings that are purpose-built smart buildings can realize those benefits from their first occupation. Older buildings that need to be retrofitted with smart building technology may never achieve

[2] 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies 2018, www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/ [accessed on April 13, 2020].

[3] 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies 2018, www.gartner.com/smarterwithgartner/5-trends-appear-on-the-gartner-hype-cycle-for-emerging-technologies-2019/ [accessed on April 13, 2020].

the full benefits of a smart building.[4,5,6] This can be attributed to a number of factors, for example, older buildings may have fewer windows, poor ventilation infrastructure, or badly situated partitioning or inner walls.

Improving Building Efficiency

Many organizations would view building efficiency as one of the major attractions of using a smart building. There are many claims made for smart building efficiency and one of the primary benefits is the reduction of energy consumption. Reducing energy consumption is attractive because of the potential cost saving and the impact on greenhouse gasses. New buildings are able to exploit cost-savings arguments when they are up for sale. Improved lighting, air conditioning, and air quality are three areas of efficiency that can have a positive effect on employees as well on the costs of owning or leasing a building. Lighting and heating efficiency are often the most obvious areas that can be improved as a cost benefit and a benefit to the occupants. Environmental factors like air quality can also have a significant effect on building occupants and their health, and even decrease tension and arguments between coworkers, as we will discuss.

Lighting

Good-quality lighting and well-controlled natural lighting are considered as a valuable tool for employees, particularly those employees who spend a long time at computer screens or monitors. Although the myth of the damage that poor lighting has caused for computer users has been

[4]Smart Buildings Magazine, https://smartbuildingsmagazine.com/ [accessed on April 13, 2020].

[5]World Construction Today, www.worldconstructiontoday.com/ [accessed on April 13, 2020].

[6]ScienceDirect, www.sciencedirect.com/science/article/pii/S0360132316303171, [accessed on April 13, 2020].

debunked,[7] good lighting can ease eye strain and improve eye health according to a study by Wout J.M. van Bommel.[8] For example, screen users tend to blink fewer times than those who are not using a screen, and this can lead to sore and irritated eyes. Users are advised to look away from the screen on a regular basis to avoid eye strain.[9] In the van Bommel study, attention is drawn to the minimizing of all these effects by working in an environment with optimal lighting.

Of course, some of the attempts to manage natural light have had comical results. There is a building in Datchet, United Kingdom, that was originally promoted by the builders as a smart building with window blinds controlled by computers and sensors on the roof. Some 20 years ago this building was occupied by the first tranche of employees who noticed that the blinds on a set of windows were either all down, all half down, or all up. The position of the blinds and the curvature of the building left some employees in bright sunlight with the majority of employees in the shade. It didn't matter what the exterior lighting values were, the blinds moved in unison. Enquiries were made and it was established that the architect designed the building to look better from outside if the blinds were all at the same level, proving that good lighting management can be easily subverted if the comfort and welfare of employees was ignored. This discovery forced changes to enable the window blinds to be moved separately by individuals, making the lighting optimal for each individual and minimizing any arguments between employees.

---

[7]Harvard Health Publishing, Harvard Medical School www.health.harvard.edu/healthbeat/safeguarding-your-sight [accessed on April 13, 2020].

[8]van Bommel, W. J. M. (2006). Non-visual biological effect of lighting and the practical meaning for lighting for work. *Applied Ergonomics*, *37*(4), 461–466. https://doi.org/10.1016/j.apergo.2006.04.009.

[9]Tribley, J., McClain, S., Karbasi, A., & Kaldenberg, J. (2011). Tips for computer vision syndrome relief and prevention. *Work*, *39*(1), 85–87. https://doi.org/10.3233/WOR-2011-1183.

Another consideration for a smart building would be the control of lighting in hazardous areas that may only occasionally need lighting. A plant room that houses environmental management and control equipment would only need lighting when a person enters the room; this can be achieved by having movement-sensitive lights triggered when a person enters. As long as the occupant of the plant room keeps moving, the light would stay on, but many of us have been in rooms with motion-sensitive lighting where the light has gone off because we are standing still, reading a meter, for example. A truly smart building would know when someone entered the room and would monitor both their activity in the room and when they leave the room. The lighting would then stay on for an optimal time and not plunge the poor occupant into darkness and force them to leap up and down to switch the light on, perhaps at a critical time for their activity. An office building that I worked in during the 1990s was an early adopter of this technology. The lights in the main body of the office were controlled by motion sensors although on permanently during the working day. The offices, however, were in the center of the building with no natural light and the office lights were solely controlled by motion sensors. Frequently, sitting at a desk, with the door closed, on the telephone the room would be plunged into darkness and I would leap up and wave my arms to get the lights on. There are also stories of people in the restroom having the lights go out.

Air Quality

Sick building syndrome (SBS) is a condition that is difficult to diagnose. It is most often reported by employees in their places of work rather than in domestic property. Possible symptoms can include headaches, blocked or runny noses, dry itchy skin, and sore eyes. Many health authorities have notes and papers on sick building syndrome[10,11] and the consensus shows that the symptoms are manifest in the building's occupants, but there may be no discernable problems with any of the building's environmental measures. There may be nothing obviously amiss with that building. There are many environmental factors that are suspected causes, and although there have been experiments in air pollution, introducing pollutants in the form of a 20-year-old uncleaned office carpet,[12] there are no clear cause and effect relationships in a "sick" building.

The symptoms of sick building syndrome are varied. It is difficult at times to distinguish symptoms of SBS from symptoms with other causes, for example, headaches may be a symptom of an occupant suffering from sick building syndrome or a symptom of eye strain from staring at a screen for too long without a break.[13] The relative causes of headaches can be tested by changing the behavior of a sufferer by reducing the amount of continuous screen time and making them take a break. If the headaches still occur, the cause may be SBS. This style of diagnosis can be used to evaluate other symptoms.

Air quality can be an important factor in the efficiency of a building and smart buildings can play a role in maintaining that quality. Dust, radon, and fungi can all contribute to a reduction in air quality. Sensors

---

[10] Finnegan, M. J., Pickering, C. A., & Burge, P. S. (1984). The sick building syndrome: prevalence studies. *Bmj*, *289*(6458), 1573–1575. https://doi.org/10.1136/bmj.289.6458.1573

[11] Burge, P. S. (2004). Sick building syndrome. *Occupational and Environmental Medicine*, *61*(2), 185–190. https://doi.org/10.1136/oem.2003.008813

[12] Wargocki, P., Wyon, D. P., Baik, Y. K., Clausen, G., & Fanger, P. O. (1999). Perceived Air Quality, Sick Building Syndrome (SBS) Symptoms and Productivity in an Office with Two Different Pollution Loads. *Indoor Air*, *9*(3), 165–179. https://doi.org/10.1111/j.1600-0668.1999.t01-1-00003.x

[13] Yan, Z., Hu, L., Chen, H., & Lu, F. (2008). Computer Vision Syndrome: A widely spreading but largely unknown epidemic among computer users. *Computers in Human Behavior*, *24*(5), 2026–2042. https://doi.org/10.1016/j.chb.2007.09.004

can detect these wherever the sensors are placed; however, placing sensors at head height can report on the quality of the air at the breathing level of the employees. Dust on the floor and in the corners of the room would not be detected unless it was disturbed. Location of sensors can be an important factor in the accuracy of air quality detection.

As can be seen from Figure 5-1 when the open plan office has been fitted out, lights and air-conditioning ducts are spaced evenly throughout the space and lighting and air conditioning were balanced in that space.

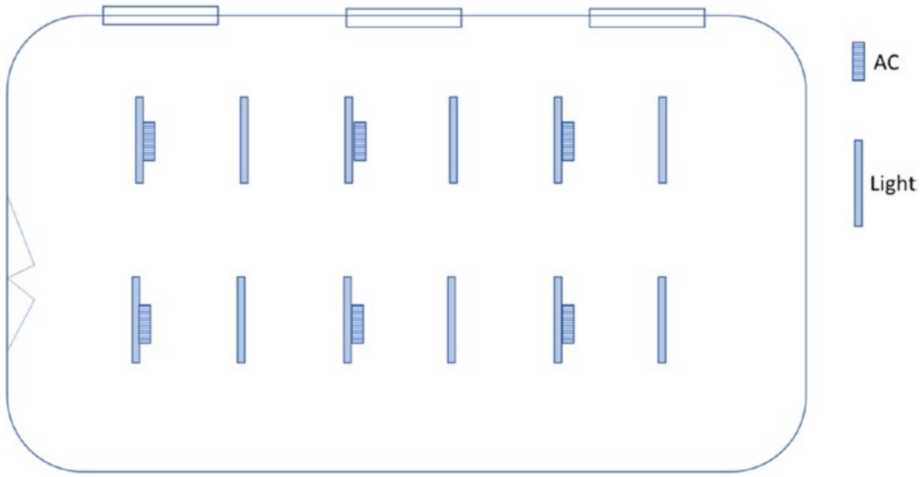

Figure 5-1. *Open plan building initial layout*

Once the new owners start building partitions for offices, the balance can be lost as shown in Figure 5-2.

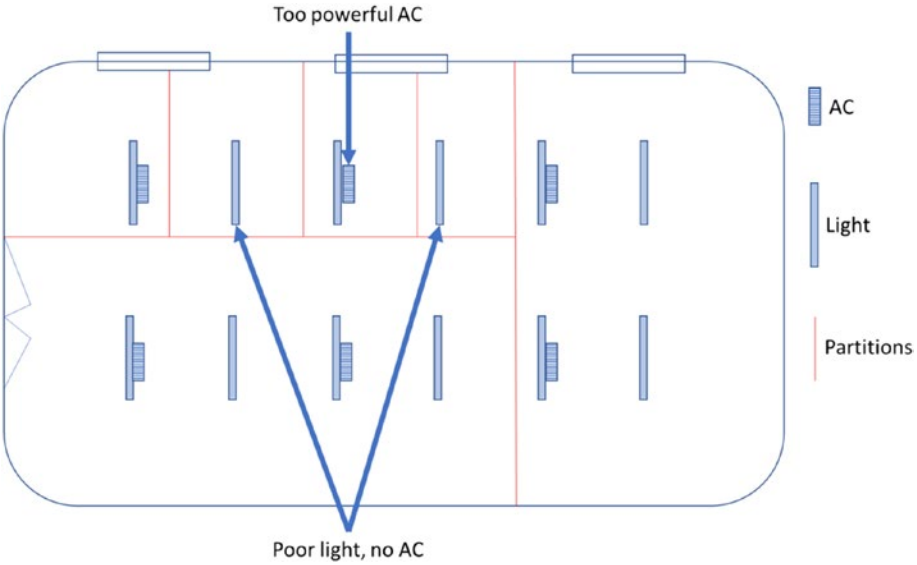

Figure 5-2. *Open plan after partitioning*

Figure 5-1 gives a well-balanced idea of where the relative utilities should be placed. If the new occupants wish to change the space by building partitions, they frequently give instructions that are not precise, for example, as long as the occupants have light it does not matter if it is at the edge of the room or in the middle This can result in the less than ideal state in Figure 5-2. The authors have experienced the effects of this; however, in one case in London, a partition was erected that split the light in half and left the control in one office. When the light was switched off, the occupant of the office without the control had to shout or go to the other office to get the light put back on. Air conditioning was often a bigger issue with one office having no air conditioning and another office having twice the throughput of heated or chilled air. On days with extreme temperatures, this could lead to some very acrimonious arguments. In a newly built smart building, the expectation is that the building is designed to have office and room partitions compatible with the requirements for optimal air conditioning, lighting, and the associated sensors to monitor those spaces. A building that is being retrofitted with partitions would be more difficult to organize, but retrofitting environmental sensors even in poorly served offices would enable some balance to be achieved. Sensors in an office that had little or no heating could indicate to environmental control staff the need for extra heating and the staff can develop a strategy based on monitoring the conditions in that room.

Temperature and Humidity

Temperature and humidity are allied to air quality and are part of the overall environment. In the paper from the Environment Protection Agency cited earlier, there is a note that the temperature in an office-style building should be kept below 23 degrees Celsius to avoid increases in SBS symptoms. In many countries the health and safety regulations specify a minimum operational temperature for an office but not a maximum. Humidity levels can also develop SBS symptoms such as itchy eyes in a dry hot environment, although this can also be attributed to screen use. Temperature and humidity levels are noted more frequently when they are at extremes. When those extreme levels are approached, occupant productivity drops. In more moderate climate where air conditioning is restricted to opening a window, it may not be possible to cool a building or a room and little or nothing of value is achieved. Smart buildings would aid productivities by eliminating extremes of temperature and humidity.

Occupant Efficiency

All of the factors mentioned earlier can contribute to occupant efficiency and in particular an occupant's productivity. You would expect that a smart building would be able to control lighting and air quality as an aid to productivity. The Environment Protection Agency cites a 1984 World Health Organization Committee report that claims potentially 30% of new and remodeled buildings worldwide may have air quality issues that can generate SBS symptoms. Air quality issues may be the result of poor design, use, or inappropriate occupant behavior. There are a number of these reports dating back to 1958 and covering pollutant effects ranging from mild symptoms such as a cough through to death. The reports cover interior pollution, household pollution, and so on.[14] It seems that problems with air quality and other environmental factors have been known for a long time. The fact that these are still considered important factors in an occupant's health may be due many reasons including cost and feasibility. It is certain that a move to a smart building where all of these factors are monitored and managed will improve a working environment. Building efficiency can be influenced by taking a more holistic view of the interior environment and that is the goal of the integrated smart building.

Most organizations may be keen to have the best working environment for staff that is also a cost-efficient solution. Some organizations may consider cost as the major factor in making a decision and conclude that it is not worth retrofitting sensors and other environmental controls in an existing building. This may be due to their financial position but without some investment organizations would lose out on a productivity bonus resulting from lowering absentee rates caused by SBS.

Increased Productivity

The health and wellness benefits should not be managed individually. The integration of the sensors into a complete picture would mean that factors are not overlooked. Fixing a failing light will return an environment to its optimal state but there should be no trade-offs; the lighting cannot be repaired at the expense of the air quality. All the benefits of smart buildings can together improve productivity and the health of the work force but only if they are integrated into a complete view of all the environmental

[14] WHO guidelines for indoor air quality: selected pollutants www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2010/who-guidelines-for-indoor-air-quality-selected-pollutants [accessed on April 13, 2020].

factors. There are a myriad of studies[15] that cite employee welfare as a critical component of productivity. These productivity advantages will be gained by improving the working environment. Employee morale is a key factor in productivity and can be affected by the health of all the work force, for example, low levels of attendance may lead to being short-staffed lowering the morale of the rest of the workers. Poor environmental control can create a poor working environment and a subsequent loss of occupant productivity.

Inefficient building design and insulation can increase energy costs and create a poor working environment. Energy costs can also be affected by the weather. The associated cost savings of retrofitting smart building technology are not so evident for older buildings but they can be measured. In some cases, there are claims that greater than 40% reduction in energy use can be achieved by fitting sensors and monitoring the CO_2, humidity, and temperature.[16] Managing energy use efficiently is another benefit of smart buildings and is allied to the productivity improvements gained from a good working environment, reducing energy use, and reducing sickness.

Resource Management

Employee skills are a major asset to any organization. There are more tangible resources and corporate assets, for example, buildings, office space office equipment, and meeting rooms. The condition of these resources has an impact on users. Imagine arriving at a cubicle and finding there is no chair or trying to use a photocopier that is out of either

[15] Hillier, D., Fewell, F., Cann, W., & Shephard, V. (2005). Wellness at work: Enhancing the quality of our working lives. *International Review of Psychiatry*, *17*(5), 419–431. https://doi.org/10.1080/09540260500238363.

[16] Smart Buildings Magazine https://smartbuildingsmagazine.com/features/how-retrofitting-smart-technology-helps-deliver-energy-efficiency [accessed on April 13, 2020].

ink or paper? The frustration and upset can be really high. The authors have too frequently been in the position of preparing a meeting room for an important meeting only to find that the vital cable connector to the audiovisual (AV) equipment is missing or incompatible with one of the presentation machines. Video conference equipment that has been left in an unworkable state by a previous occupant is another problem that causes anxiety in the organizer of a meeting. Locating employees in an unfamiliar floor or building can also take time and repeated phone calls; room booking can also have its problems. Using the organization's resources can be frustrating at times and counterproductive. In a smart building it would be possible to access the meeting room information before booking to see the condition of the AV equipment. You can discover the location of a room if you are unfamiliar with the building and the location of a member of staff if you don't know where they are working that day. All of these solutions can come as part of a smart building management system.

Maintenance has been mentioned several times in this section of the chapter, and in a working building, wear and tear can occur. Preventive maintenance is a potential feature of smart buildings. Data analysis can be used to predict failures in building components, for example, a door that has been opened frequently and is getting close to the mean point of failure for the hinges. iSmart building systems could identify the door and its potential for failure. The smart building system can notify maintenance staff of the need to examine, order spare parts, and plan repairs before failures disrupt the use of the space. What would happen if a door hinge fails when you are in the room? Are you trapped? Would the door injure someone if the hinge failed? The consequences of a failure could be avoided if the potential failure of the hinge was known in time for preventive maintenance.

Not only do smart buildings manage and maintain the building environment, smart buildings are also capable of integrating environmental, operational, and resource data and producing a more efficient experience for the occupants. The keyword in the last sentence is integrating.

As mentioned, employees can check on the meeting room and equipment status before they decide to book it. If the employee's diary system is integrated with the room booking system, the booking system can take information about dates, time, and attendees without the employee having to switch from system to system. The booking system now knows the details about a meeting and can contact the maintenance department with the equipment failure notice and a request for repairs. There are many other tasks that can be accomplished in a smart building management system, and the possibilities are huge; however, they still need all of the internal and external controls integrating with each other. The management system may also require techniques such as data fusion that are more commonly associated with robotics and autonomous vehicles.

Smart Building Example

The authors were able to visit a building that has an integrated smart building management system. Tieto's Empathic Building Technology[17] has been installed at the headquarters of Tieto Finland Oy at Keilalahdentie 2-4, 02150 Espoo, Finland. This is a purpose-built smart building in Finland. Tieto's objectives, stated on their website,[13] are

- Boost happiness
- Increase performance
- Better experiences

[17]Tieto Empathic Building, `www.tieto.com/en/what-we-do/data-and-ai/tieto-empathic-building/` [accessed on April 13, 2020].

The benefits of smart buildings have already been laid out in the earlier part of this chapter and Tieto plans to take advantage of these benefits. An emphasis on wellness has the potential to be one of the most important results from smart building technology. Tieto tells us on their website that smart buildings will help develop a "more motivated and agile workforce with a better workflow."

The headquarters, an impressive structure of steel and glass, is a functioning smart building and a demonstrator for their integration products. Environmental and operational information sources are integrated to create and manage a more complete data set that covers a wide variety of factors from heating and lighting to room booking. It is the integration of all the data from all of the information sources that makes the building smart. As we have noted earlier, there are many hard objectives for smart buildings based around cost management, energy management, and building efficiency. We have also described some of the effects of a smart building on the occupants including alleviating symptoms of sick building syndrome, increasing employee well-being, and removing irritations generated by various environmental factors. When we entered the Tieto building, there was a large display in the lobby announcing events and visitors. We went to the reception desk and they were able to tell us how far away from reception our host was and how long it will take for them to meet us; that gave us the opportunity to take a coffee and relax. As soon as we registered with reception, our host was notified that we had arrived and where we were. Wearing a visitor's badge to gain access to the elevators ensured that we had a friction free visit and you could give full attention to the meeting. During the meeting, we viewed the smart building data and we were shown information that helps Tieto optimize their use of space. If an area or room was not being fully utilized, they would see it in the data. From there Tieto can plan to repurpose the space or change the configuration of that part of the building. All of this information was readily available on their application, and it was fascinating to see a display that could show the meeting room we were in, who was in there, what the meeting topic was, and what equipment was being used. This was a very interesting meeting but most importantly we felt very relaxed. It may not seem a big advantage but the knowledge that we could leave the meeting to visit a lavatory or to get a drink of water without having to be escorted or borrowing an identity badge for access made it easier to concentrate on the meeting. All the doors responded automatically according to the security level of the visitor's badge and what security zone we were trying to enter. At the end of the meeting, we needed no escort to get to the elevators and we signed out of the building by handing over our visitors' badges. We remarked on the ease of attending a meeting where the meeting was the focus, not can someone tell us if we are going to wait long for our host.

Smart buildings and the associated technology are being marketed and developed in increasing numbers. Increased initial costs of a smart building can be offset by long-term cost saving from better energy management and staff well-being. All of these improvements are based on using integration of data from environmental and operational technologies that deliver the benefits that we have described.

In the future many people will benefit from the implementation of smart buildings. Some problems in older buildings will need to be addressed by using integrated approaches to a building environment. Sick building syndrome is a good example of an issue for occupants and using smart buildings environmental controls will benefit the building occupants. Benefits for organizations including increased productivity will provide productivity and improvements in morale. In addition to the environmental benefits, the integration of internal systems such as booking, diary, and other supporting applications can increase the interaction between the occupant, the building, and the organization as a whole. The building then becomes autonomous, gathering data and making decisions based on data from occupants and the building environment that will improve all aspects of the physical organization. With these advances, the smart building, its environment, and occupants can then be thought of as a robot, but without arms.

Autonomous Vehicles

There is a large body of work covering the technical, human interaction, and impact of autonomous vehicles on society. Autonomous vehicles frequently attract attention from journalists who sensationalize the successes of this technology as well as the failures. There are many articles and papers focused on this technology. The articles record the strides in safety, the advantages, disadvantages, and frequently the timescale for widespread adoption of autonomous vehicles.[18,19,20] This part of the chapter will review autonomous vehicles from the standpoint of the impact of personal and commercial autonomous solutions on employment and in particular the future of work. It is worthy of note that the autonomous vehicle domain is moving at such a speed that a major advance can be made in a matter of months rather than years. The societal impact will become more obvious once autonomous vehicle adoption is as widespread as the adoption of electric and hybrid cars are in 2020.

Fully autonomous vehicles are capable of making decisions based on information from a wide variety of sources, including infrared sensors, LIDAR, radar, GPS, cameras, mapping of the environment, input from street furniture, and other sources of data. Autonomous vehicles are capable of driving from one place to another without human intervention and can communicate and collaborate with the outside world.[21] They are

[18] Marvin, B. (2019, March 1). Which Self-Driving Cars Put in the Most Fully Autonomous Miles? Retrieved February 27, 2020, from `www.pcmag.com/news/which-self-driving-cars-put-in-the-most-fully-autonomous-miles`

[19] Davies, A. (2019, April 22). Are We There Yet? A Reality Check on Self-Driving Cars. Retrieved February 27, 2020, from `www.wired.com/story/future-of-transportation-self-driving-cars-reality-check/`

[20] Kaan Ozbay, Xuegang (Jeff) Ban & C. Y. David Yang (2018). Developments in connected and automated vehicles, Journal of Intelligent Transportation Systems, 22:3, 187-189, DOI: `https://doi.org/10.1080/15472450.2018.1466407`

[21] Young, M. (2020, February 2). 43 Examples of Autonomous Vehicles. Retrieved April 12, 2020, from `www.trendhunter.com/slideshow/autonomous-vehicles`

independent from the requirement to have a human guiding them. The information source "Which?[22]" has a table of levels of vehicle automation based on similar levels in a National Highway Traffic Safety Administration (NHTSA) document that can give an insight into the current positions of solutions for autonomous vehicles as vehicles. The NHSTA document also outlines safety and regulation options.[23]

Table 5-1 indicates an approximate status of automated vehicles although different manufacturers may claim a position higher up the spectrum of automation. An article in the *New York Times* indicates that many autonomous vehicle projects from manufacturers are not progressing as fast as first anticipated.[24] There are several claims made for the value of automation from improved safety to real assistance in a medical emergency, for instance, in an article in *The Guardian*.[25] In the case cited the driver was having a medical emergency, switched on the autopilot, and got driven near enough to the hospital to be able to take over for a few minutes and guide the car to the ER entrance. It can be claimed that this vehicle is approximately at level 3, Conditional Assistance, because the driver still had to touch the steering wheel regularly to prevent the autopilot driving to the verge of the road and stop. Sadly, there are also a number of cases, noted in the same article, that

[22]Harding, J. (2018). Driverless Cars: What Are Autonomous Vehicles? Retrieved February 27, 2020, from www.which.co.uk/reviews/new-and-used-cars/article/driverless-cars-what-are-autonomous-vehicles

[23]National Highway Traffic Safety Admin (NHTSA) (2013). US Department of Transportation, preliminary statement of policy concerning automated vehicles. *NHTSA preliminary statement*. www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf

[24]Boudette, N. (July 17, 2019). Despite High Hopes, Self-Driving Cars Are "Way in the Future." Retrieved April 11, 2020, from www.nytimes.com/2019/07/17/business/self-driving-autonomous-cars.html

[25]Tesla's autopilot helps get man to the hospital during medical emergency | The Guardian. Retrieved February 28, 2020, from www.theguardian.com/technology/2016/aug/08/tesla-model-x-missouri-medical-emergency

claim accidents and even death that are also partially attributable to the driver using autonomous driving technology. The reality is that automated vehicles are still in the early stages of development and deployment. Other related areas still in the early stages of development are the legal, insurance, and regulatory domains. These may be critical to the future of work; however, their potential impact is not well understood in the employment domain.

Table 5-1. *Levels of Vehicle Automation*

Level	Meaning	Description	Status
0	No Automation	A majority of vehicles on the road are at this level.	Ubiquitous.
1	Driver Assist	Some autonomy but under passenger control, e.g., cruise control or lane assist.	Been in place for more than 20 years.
2	Partial Assistance	Driver is aware and capable of taking full control. Increased autonomy when engaged, e.g., traffic-aware cruise control.	Still fairly new.
3	Conditional Assistance	This requires a driver and the vehicle can carry out many tasks, but the driver must be prepared to take over if necessary.	Some vehicles can match this.
4	High Automation	Once the vehicle is in an appropriate environment, the driver can switch on the automation and relax, e.g., on a freeway.	Not possible today except in experimental areas.
5	Full Automation	Does not need any external control, steering wheel, or pedals. No human control or driver necessary.	Not possible today except in experimental areas.

Challenges and Triumphs

While there are still a number of challenges to autonomous vehicles becoming a normal mode of transport, there have also been a number of victories. We will not dwell on the potential insurance and legal issues should the person have an accident while driving to the hospital.

Maps of metropolitan areas held in the vehicles internal data store can help guide an autonomous vehicle, taking notice of regular obstacles, red lights, and street furniture. Some maps can be updated with more temporary obstructions, parked vehicles, broken-down vehicles, or garbage on the road. This would be enabled by autonomous vehicles communicating and sharing information. This has the potential to increase safety based on a better awareness of driving conditions. Mapping technology allied with sensor input from the vehicle and the environment is one of the cases where data fusion (see Chapter 7, "Robots in Society") becomes important. Although a fully autonomous vehicle has little need for a driver, the vehicle itself has to blend different data into a complete view of its environment to avoid accidents and drive successfully.

The ethical issues that are raised by driverless level 5 fully autonomous vehicles are no less challenging than successful navigation of the environment. Most people would agree that if there is a pedestrian in the roadway, the vehicle should avoid the pedestrian if possible, and if not, the vehicle should stop. So far, so obvious. An interesting question is raised if the pedestrian in the roadway is a criminal carjacker, what is the decision in this event? Most people would either take avoiding action but would draw the line at running the pedestrian over. A further dilemma is based on the well-known trolley problem.[26] In autonomous vehicle terms the problem has been outlined in a paper called "The social dilemma of

[26]Thomson, J. J. (1985). The Trolley Problem. *The Yale Law Journal*, *94*(6), 1395–3. https://doi.org/10.2307/796133

autonomous vehicles.[27]" If an autonomous vehicle detects six pedestrians in the road and it cannot stop in time, should it swerve to avoid them even if that swerve would kill a single pedestrian on the side of the road? In a refinement, should the vehicle swerve into the wall at the side of the road, saving all the pedestrians but killing the occupants of the vehicle? These moral questions are difficult to resolve, even if the driver has a coin to toss. Studies have shown that in the case of killing pedestrians many people would take the reduced casualty option[28] but would be more reluctant to drive into a wall saving the pedestrians at the cost of their own life.

The real triumphs will come in the future with level 5 automation. Safety improvements, reduction in parking requirements, and reduced numbers of vehicles on the road will improve pollution, congestion, and general well-being. However, safety may also be considered a major block to declaring a vehicle fully autonomous. Road safety is something that children learn about at an early age, usually focused on creating awareness to traffic and avoiding accidents. As the children grow into young adults and learn to drive, they are taught safety from the point of view of the driver and are expected to understand the risks from the point of view of pedestrians.

Mobile phone habits have changed this a little; the expansion in the number of people with mobile phones has led to an expansion in the number of pedestrians who suffer from cognitive distraction.[29] This leads to an increase in accidents since there are a number of distractions

[27]Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016, June 24). The social dilemma of autonomous vehicles. Retrieved February 28, 2020, from https://arxiv.org/pdf/1510.03346.pdf

[28]Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016, June 24). The social dilemma of autonomous vehicles. Retrieved February 28, 2020, from https://arxiv.org/pdf/1510.03346.pdf

[29]Nasar, J., & Troyer, D. (2013). Pedestrian injuries due to mobile phone use in public places. *Accident; Analysis and Prevention*, *57C*, 91–95. https://doi.org/10.1016/j.aap.2013.03.021

for drivers already and having pedestrians distracted leads to situation where a moment's lack of attention compounded by pedestrians gazing at their phones can lead to more accidents. Tiredness or lack of attention particularly on long journeys would also be eliminated and result in safer roads. Tiredness that may cause lane drifting is already addressed by lane management technology installed in some vehicles with level 2 automation.

There is an anticipated reduction in the number of vehicles on the roads based partly on changes to driving habits and traffic. Improved traffic flow will result from fully automated vehicles integrating their internal data and fusing it with data from street furniture, traffic lights, traffic signs, and other vehicles. This fused data can be used to allow vehicles to travel closer than is advisable in a regular vehicle. Fused data from other autonomous vehicles nearby, in front, behind, and at the side of a vehicle can be collected. Internal data on braking, steering, or lane changes can be fused with the external data and can be managed at computer speed eliminating human error and variable human reaction times. Vehicles will have enough information to join traffic flows without disruption to other vehicles and still maintain the traffic density.

Impact on Society

Personal vehicles have a special place in many people's hearts. An elderly relative of one of the authors explained their need to own and drive a car, even in their 80s, in one word, "independence." On questioning they said that without the car they are dependent on others for shopping, trips out, and going to see other relatives. Extending the travel time by using infrequent public transport does not compare well to the independence of driving a car. Talking about a relative giving up a car is frequently a difficult topic to discuss due to the level of emotion that is generated. The use of fully automated vehicles has the potential to resolve this issue, as the increasing age and driving capability of a user will no longer be a barrier to travel. Passengers in an autonomous vehicle are just passengers in a vehicle that does not need a human driver. Elderly drivers whose eyesight or other problems could ban them from the steering wheel can still be mobile. Impairment of the faculties of a passenger will no longer be considered a barrier to travel. The highest level of automation will remove some if not all the responsibilities from a vehicle occupant.

When looking at how most users will interact with autonomous vehicles, we should consider a number of scenarios. The first scenario is a world where nothing changes in the behavior of the vehicle user. They own their own vehicle and use it in the same way that they would use a conventional vehicle. Scenario one is least likely in the long term due to financial and societal issues outlined in the following “Private Vehicles” section. Scenario two is where the new technology results in a new attitude. A user may decide that they can work in the autonomous vehicle and have no need to stay late at work. Even the most optimistic commuter cannot accomplish much in a crowded and cramped commuter train, but in the relative peace of an autonomous vehicle, they may be able to work successfully. They would include their commuting time in their working time. They would continue to own their vehicle. Scenario three could be the new technology would result in a change in attitude to both working and ownership. Using a shared pool of autonomous vehicles will reduce costs and will reduce congestion by having fewer cars on the road. After establishing the value of an autonomous vehicle to passengers in all scenarios, there will be a need to establish requirements for a dedicated personal vehicle. In this chapter we have chosen the third scenario as our focus since it will be the most radical. It is also possible that there is a mix of scenarios. A scenario that suggests no behavioral change is unlikely to be happen in the long term.

Private Vehicles

Owning an autonomous vehicle as a private individual may be prohibitively expensive. The running costs will remain at least as high as for conventional vehicles. Additional expenses such as updated software and data storage as well as power and maintenance requirements will likely increase costs. Just keeping a conventional vehicle on the driveway costs money and private car utilization is at approximately a mere 4%. In urban areas there is an attraction to sharing vehicles, as long as the shared vehicle is available almost on demand. Shared autonomous vehicles will further lower the costs of commuting and shopping. Strategists and analysts such as UBS[30] speculate that future autonomous vehicle-based car sharing can replace 25 private cars by 1 shared vehicle. Ride hailing can replace an estimated 5 to 10 cars. Ride sharing is already available in large urban areas with Uber. Vehicle sharing using the Zipcar model is currently available in large urban areas. Zipcar is probably the closest analogy to autonomous vehicle sharing with the main difference that an autonomous vehicle can drive itself from its park to the customer without any human in the car. Zipcars require you to travel to their parking place. The reduction in number of vehicles, lower vehicle emissions, and lower environmental impact will make a compelling case for individual and governmental support for ride share and ride hail.

Commuting in this future will be different. A driverless car arrives at home, the commuter enters and is taken to the train station or to the office and the vehicle leaves for a new assignment, and other vehicles will do the school run and shopping trip all without the need to large areas of land being dedicated to car parks.

[30] UBS: Ganter, R., Berrisford, C., Dennean, K., & Dessloch, S. (2019). Smart mobility, 1–31. Accessed on February 28, 2020, from www.ubs.com/global/en/wealth-management/chief-investment-office/market-insights/digital-disruptions/2017/smart-mobility.html

Figure 5-3 shows a potential model for using ride share and ride hail with the scheduled regular requirements being met from a pool of vehicles, with the vehicle arriving at a set time. If there are requirements that are outside the contracted ride share times, there is also the option to use a ride hail vehicle that will have a lead time to arrive but may have availability problems. A cautionary note is that this model will work best in an urban environment; rural areas have their own challenges due to low population density, although not all urban areas are equal in support. The preponderance of autonomous vehicle development and testing in the Silicon Valley area has resulted in highly detailed road mapping. Other urban areas both in the United States and other countries have fewer development and test areas and maps are subsequently less detailed.

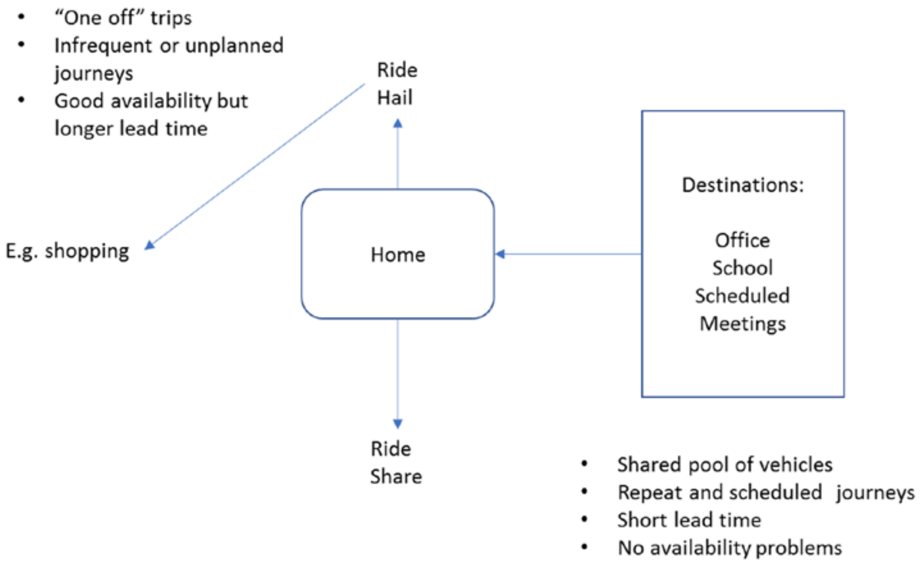

Figure 5-3. *Ride hail/share strategy*

The impact of private autonomous vehicles on employment is likely to be changes in individual working practices. Travel in a private autonomous vehicle would facilitate work during travel time with no detriment to the vehicle safety. Passengers can read, type, and call with colleagues, even attending conference calls over the in-car Internet, ensuring that a long commute is no longer wasted time. A trouble-free commute may be jeopardized should the occupant suffer from motion sickness. Reading is frequently cited as a cause of motion sickness[31] and working in an automated vehicle may generate this unwanted side effect.

Driverless vehicles will mean fewer human drivers. Fewer cars will be sold and the many companies that support the manufacture of personal vehicles will also be under threat from fewer cars being built. Employment in the car industry has been disrupted over time and will continue to be disrupted as autonomous vehicles become common. Even the service centers will employ more computer engineers than oil changers. Disruption of the personal car domain will not be as immediate and painful as the disruption in the transport and delivery industries.

Commercial Vehicles

Self-driving autonomous vehicles will be the cause of a massive disruption in the social and economic life of developed and developing nations for years to come. Freight carrying and passenger carrying autonomous vehicles will result in new working practices and fewer employees in many cases. Licensed taxi drivers in most cities are already being squeezed out of work by the working practices and business model of companies like Uber and Lyft. Even the new working practices of these two companies with sophisticated mapping, ride hailing, and pricing will not be able to compete with an autonomous taxi that needs no driver and needs no break between passenger. An autonomous taxi is expected to be safer and more cost effective for both the ride hailing company and the customer, removing the only intermediary left, the driver. All will not be doom and gloom for drivers. Luxury cars may still use human drivers as a status

[31] Hain, T. (2003, February 17). Why does reading in a moving car cause motion sickness? Scientific American. Retrieved March 3, 2020, from www.scientificamerican.com/article/why-does-reading-in-a-mov/

symbol. There is also a potential demand for delivery drivers for first and last mile delivery. If you consider a supply chain that looks like Figure 5-4, goods can be delivered a short distance to a warehouse where they are loaded on a freight autonomous vehicle. City to city highways are much easier for autonomous vehicles to navigate than small town and rural roads, with fewer dynamic changes or obstructions to the roadway.

Goods will need to be moved from the manufacturer to the distribution center, in this case the First Mile Warehouse. Movement can be done with driverless or driven vehicles. At the warehouse the goods are loaded on to autonomous vehicles that only need simplified mapping to get them to the highways.

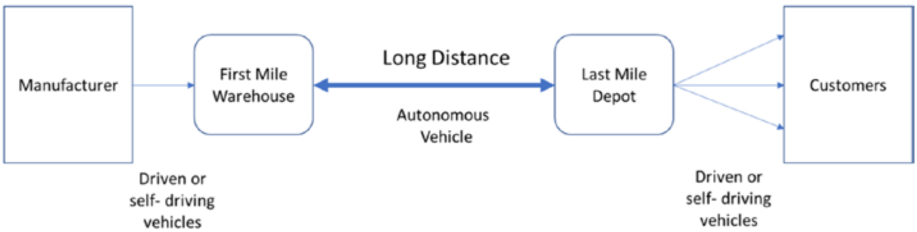

Figure 5-4. *Last mile delivery strategy*

This vehicle will then travel the longer distance without a driver, dropping goods at so-called “last mile” depots. The goods can be unloaded by robots and loaded onto smaller vehicles that may have a supervising driver. The importance of the last mile is that this is often the highest cost of the delivery particularly in pollution terms. Using electric and supervised self-driving, this can potentially offer an increased demand for local delivery drivers. Local delivery can remove many of the health issues of the long-distance driver. The increased demand for local delivery may reduce the impact of automation on the working population of truck drivers.

The strategy of replacement of truck drivers by automated vehicles is progressing with many of the leading truck makers developing vehicles that can drive themselves. The trucking industry is ripe for automation.

There are not enough drivers to meet current demand and those that are in the industry have serious ailments from the physical nature of the work, the mental stress, and frequent long periods away from home. The replacement of drivers is only now becoming visible to the truckers themselves. Some commentators are anticipating labor disputes and even riots as more and more trucker replacements become visible.[32] Other commentators see a need for retraining of truckers who lose their driving jobs into new industries or the more technical aspects of automated vehicles.[33] Autonomous vehicles will have many sensors, radar, LIDAR, and other technology, onboard computers, and communications; all of which will be severely tested by the rigors of driving for 24 or 48 hours with no stops. Maintenance of autonomous trucks will be more complex than brakes, lights, and the drive train.

Elsewhere in this book we have commented that automation of repetitive, often low-paid work results in the creation of different new jobs, and this may be the case with commercial vehicles. Trucks are likely to become so automated in the future that they can also be called robots without arms. They can be carrying out valuable work while freeing humans to do more creative work.

Summary and Conclusion

In the introduction we noted that smart buildings and autonomous vehicles exhibit the attributes of a robot: self-directing, decision-making, data gathering, responding to collaboration, and so on. However, the main

[32]Yang, A. (March 4, 2020). Self-Driving Vehicles: What Will Happen to Truck Drivers? Evonomics. Retrieved March 4, 2020, from https://evonomics.com/what-will-happen-to-truck-drivers-ask-factory-workers-andrew-yang/

[33]Samuels, A. (February 27, 2017). The Automation of Trucking. The Atlantic. Retrieved March 4, 2020, from www.theatlantic.com/business/archive/2017/02/when-robots-take-bad-jobs/517953/

differentiator is that either of these special cases is unlikely to manipulate objects. It is possible to reaffirm that smart buildings and autonomous vehicles can be called robots without arms.

Smart buildings may not have a big disruptive effect on employment and the future of work. These can be a considered as a more efficient way of housing workers rather than a robot doing work. Workers in smart buildings will, however, respond well to an optimized environment and are generally more productive in a smart building setting. An additional bonus for employers will be the improved health of the employees that work in the smart building. A causal relationship between poor built environments and the symptoms of sick building syndrome may be difficult to establish, but there is evidence that the costs of improving an employee's environment are easily offset by increased productivity and fewer sick days. A section on the economic consequences of sick building syndrome in Burge's paper[34] supports this assertion.

Large amounts of data from sensors both inside and outside the building can be collected and integrated. This will allow better and more comprehensive data analysis based on that data. In the future it will be possible for an employee to be recognized when they arrive at their work with an optimal car parking space being allocated and a notice of any changes to a meeting that they are attending can be sent to their car or phone. Efficient use of office or warehouse space will be the result of analysis showing the space utilization and the probability of a space being under used or overused to maximize the working area.

The future of work over the next 10 years will see employees and buildings integrated to create a pleasant and productive work environment without those frustrations and sickness that can create stress and anxiety for the people delivering value in the workplace.

[34]Burge, P. S. (2004). Sick building syndrome. *Occupational and Environmental Medicine*, *61*(2), 185–190. https://doi.org/10.1136/oem.2003.008813

Autonomous private vehicles are likely to have a tangible effect on the future of work resulting from their effect on the employment of freight-based drivers of commercial vehicles. The use of autonomous vehicle technology for both commercial and private vehicles is similar in its goals and objectives. The description of private vehicle functionality closely resembles that of the commercial vehicles. Private autonomous vehicles will change some working practices, with commuters being able to live further away from the office or use their travel time for effective working. Nondrivers will also become less dependent on others for mobility by having access to driverless cars. Passengers who are young, elderly, or unable to drive will no longer have to rely on a third party to take them where they need to be. The removal of an obstacle to travel for work or amusement will potentially generate a more self-reliant society.

Autonomous commercial vehicles will have a serious effect on the employment of truckers and freight delivery drivers by replacing them. Once these vehicles are common, the numbers of drivers will reduce massively. In the shorter term, it is likely that long-distance truckers in the United States will still be needed to supervise vehicles or conduct pods of trucks, but these will eventually be replaced by fully autonomous vehicles. Prospects for low-paid repetitive jobs will be poor in the automated future; however, vehicle maintenance will become more technologically oriented. Maintenance will not just be about the brakes, lights, and drive chain of a truck but the sophisticated electronics that will be installed, from sensors to computers and communication devices. Trucks are likely to get external data about weather, obstructions, and jams from the data that they and other trucks gather and store in the cloud for analysis. All of this will offer an opportunity for truckers to retrain in technology and other jobs. New delivery models will also contribute to reemployment of truckers as last mile solutions will need an increase in delivery drivers.

In both these cases data is of great importance. Smart buildings with integrated software will store data for analysis and automated action. Autonomous vehicles will need to access data about external conditions as well as vehicle telemetry. Analysis of this data will increase the ability of managers of both smart buildings and autonomous vehicles to institute preventative maintenance and reduce the downtime of either smart buildings or autonomous vehicles. Smart buildings and autonomous vehicles will require more sophisticated maintenance personnel.

PART III

Making Sense for Robots and Society

CHAPTER 6

Robots in a World of Data

Creating Human-Robot Synergy

Data fusion is considered an important component for many of the new technology initiatives that are being developed. We define data fusion in the next section of the chapter. Data fusion is an enabling technology in many domains, but currently it has been brought to prominence by the development of autonomous vehicles, intelligent automation, and collaborating autonomous mobile robots.

The importance of data fusion in robotics is attributed to a need for a common understanding of all relevant data points and decisions based on those data points. When a collaborative robot is holding a piece of wood, data is being collected by sensors on the whole of the robot's manipulator. The data can describe the parts of the manipulator holding the wood, the pressure that is being exerted on the wood, and the direction of motion of the manipulator. All of this and additional data are not immediately understood by the human collaborator but has to be interpreted programmatically along with other data that may come from elsewhere either on the robot or natural language instructions to the robot or identical data from other. All of these different types of data, collected by robots and humans, must be merged together to create a model of the

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_6

environment the collaboration is taking place in. Collaborative robotics needs a reliable method of communication and decision-making that is understood by all parties. Autonomous vehicles are leading this initiative using the combination of different sensors inside and outside the vehicles. Just travelling down an empty road requires a huge amount of data from sensors including LIDAR, radar, video, and ultrasonic. The data from these sensors is used to calculate the distance and shape of objects, how fast they and the vehicle is moving and in what direction. All of this data informs a series of decisions on direction, speed, and safety of the vehicle and its surroundings. Without the fusion of this data, the vehicle may not be able to move safely.

This chapter will examine data fusion in the domain of collaborative robotics and in particular a research project led by Professor Moncef Gabbouj of Tampere Technical University, Tampere, Finland. In conjunction with Professor Gabbouj, we feel that successful data fusion is a critical technology in the future development of collaborative robotics. There will be additional input from the ENACT research project, in the domain of healthcare. The ENACT objectives are based on applied research and will shed light on some of the architectural barriers to a comprehensive and future solution.

The main difference between the type of approach of an autonomous vehicle and a collaborative robot is the different interactions between the human actors and the vehicle/robot. A collaborative robot needs to share in the decision-making that leads to a complete task, but a fully autonomous vehicle is designed to have total control of the decision-making process, requiring no human intervention once the goals of the journey have been set with the passengers being no more interactive than cargo. The main task of data fusion in autonomous vehicles is to maintain local mapping to enable a robot to locate itself in the learned environment and its position relative to other robots.

Data Fusion Definitions

In a “A Review of Data Fusion Techniques,”[1] Federico Castanado notes that information fusion and data fusion are often considered synonymous.

The article also notes that in some situations data fusion and information fusion can describe different effects. For example, data fusion handles raw unprocessed data and information fusion can be described as the fusion of processed data. There are many definitions of data fusion but the most generally accepted definition of data fusion was generated by the Joint Directors of Laboratories (JDL). Akiwowo and Eftekhari refer to the JDL model developed in the mid-1980s and JDL is still cited in many papers.

> *Data Fusion: a multilevel multifaceted process dealing with the automatic detection, association, correlation, estimation and combination of data and information from single and multiple sources to achieve refined position and identity assessments of situations and threats and their significance.*[2,3]

There are other definitions from organizations such as the Institute of Electrical and Electronics Engineers (IEEE). Perhaps the most succinct is that defined by Steinberg in 1999. This follows on from a definition by Hall and Linas[4] and is a refinement of JDL.

[1]A Review of Data Fusion Techniques, Federico Castanedo, Hindawi Scientific World Journal 2013.

[2]JDL, *Data Fusion Lexicon. Technical Panel For C3*, F.E. White, San Diego Calif 1991.

[3]Feature-based detection using Bayesian data fusion, Akiwowo, Ayodeji, Eftekhari, Mahroo, 2013/12/01, International Journal of Image and Data Fusion.

[4]D. L. Hall and J. Llinas, “An introduction to multisensor data fusion,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6–23, 1997.

> *Data Fusion is the process of combining data to refine state estimates and predictions.*[5]

With all these definitions and possible conflicts between data fusion and information fusion, we are settling on data fusion as a more general definition than information fusion in this chapter. The JDL model does more than have a snappy definition of data fusion. The model additionally has several distinct layers that define the status of a data fusion exercise. This has been developed by the JDL Information Group as illustrated in Table 6-1. This layered structure of the modeled data fusion status is not without its critics. A critical problem with layered models is that they imply some form of order to the different layers, although this does not need to be the case. A layered model can also imply that there are no humans in the loop, but there may be frequent interventions needed when something in the process is not understood or there is an error. Despite these limitations the model is excellent for visualizing the process.

Table 6-1. *JDL Data Fusion Model*

Level	Description
0	Source preprocessing/subject assignment
1	Object assessment
2	Situation assessment
3	Impact assessment (or threat refinement)
4	Process refinement
5	User refinement or cognitive refinement

[5]Revisions to the JDL data fusion model. Steinberg, Bowman, White Aerosense99, International Society for Optics and Photonics 1999.

Humans: The Data Fusion Mavens

Mavens are a trusted expert in a particular field—in this case, the field of data fusion. If it seems as though we are discussing data fusion capabilities as something new, nothing can be further than the truth. Back in the mists of time, early hominids were able to combine data from a number of senses to help them survive. They fused data from vision, taste, and touch to tell if a fruit was ripe and if it was safe to eat. A red fruit looks ripe; if it is soft to the touch, it is even more likely to be ripe. If other animals are eating it, then it is not likely to be poisonous; however, if it is bitter to the taste, then it may not be something to eat unless in extremis. Data fusion is not new. Data fusion is a part of our everyday life as a human animal. Once we are past the immobile baby stage, we become increasingly sophisticated in our use of data fusion. Hearing, vision, touch, taste, and indeed all our sense are providing data to the brain and being used to create a complex model of the world we live in.

Humans don't integrate data into a single truth. Our approach to data fusion is to evaluate and discard data that is of no interest at the time. This is an automated process. Humans will view the whole scene in front of them and then ignore irrelevant data. We can't concentrate on everything at once; we have to critically evaluate the data we are receiving and ignore data of little or no interest. If we don't have accurate data that covers all the scenarios, we use personal or shared experience that can give an approximation that may be good enough to take valid decisions. For example, if you are driving a car in the rain, you may not know more about the state of the road than it is wet. If you are driving on a familiar road, you may feel comfortable driving a little faster or slower based on your memory of the road. If the road is an unknown, you may have to guess on the road conditions, and if you can see the countryside, you may notice a hilly environment that may be liable to flooding and slow down accordingly. These inferences, based on personal experience, and the fusing of data on the condition of the vehicle, the weather, and the topography may give a good enough overall view that allows you to continue on your journey or pull over and wait until the weather improves. This ability is not something that can easily be programmed into robots, and this chapter will look at the attempts to create a robot that can take this type of decision based on the fusion of data from many sensors and the development of "experience" that will improve decisions on the same way that the Human Data Fusion Maven can.

First Steps in Data Fusion: Structured Digital Data

It is common when discussing sensors and actuators to think in terms of structured digital data, or signals that can be interpreted as digital data. Many sensors have descriptions of the data they provide and the structure of that data. Any confusion about the type of data to be processed may lead to simplifications and suppositions creating a false impression of the challenges of data fusion. For example, in a sensor-based environment where sensor signals are sent in a "standard" structured digital representation, they can be handled as much of the nonsensor digital data processed in an environment. Sensor data can be used as input to various storage or processing systems and analyzed. The main challenges in using structured digital data will always be those of scope, timing, and lack of standards. It is not enough to pile the data into a data store and link the data together for analysis. Sensors will have different sampling rates, timing measures, and most likely be from different parts of a highly complex environment. When a robot hand is grasping a glass vessel like a jar, there will be sensors in the robot hand and arm that are generating data although the data format and sampling rates will most likely be different. Some data will come from pressure sensors that can tell how much pressure the hand is exerting. Other data from pressure sensors can be used to establish the coefficient of friction between the glass and the hand. These data, once fused together, may result in a decision by the robot to increase the pressure of the hand. A glass jar is a delicate object and the robot using data from the hand has to calculate the optimum amount of pressure to hold the jar or place the jar down before it slips from the grasp of the robot or the robot crushes the jar. Similarly, if the robot arm is in motion, with the hand holding the glass jar, there is more data collected, analyzed, and fused with the data from the hand so that height direction of travel and weight data can be added to the other fused data to give a more comprehensive model of the work being done, changing the level of decision from is the robot hand going to crush the jar to the robot can complete its task to carry the glass without crushing it. Different sensors generate different data, but similar sensors measuring, for example, moisture, may also have data differences. If a moisture sensor is being used in a field, the sampling rate and data generated will be different to that of a moisture sensor that is measuring a grain silo and different again from a sensor measuring moisture in a hay store. All the sensors will be similar and may even be the same model but a field moisture sensor may be sampling hourly, the grain store daily, and the hay store weekly. The data generated will be the same, but data fusion will be needed to ensure the accuracy of the moisture model of the farm. Data fusion is needed to enable risk and decision models to be created and, if needed, remedial action to be taken.

Autonomous vehicles have a different layer of complexity. A wide variety of monitoring data is important for safety and effectiveness in a poorly unstructured environment. Some of this data can be offline, for example, traffic flow monitors. Other data like that provided by onboard Doppler radar provides point information about the velocity of objects at a distance but is not able to provide geospatial data that would need a different sensor that can give coordinates or an address. All of these different types of data can be converted and held in a database ready for processing into data for decision-making data input, fused or not.

Decision-making using structured data is not error-free. There have been several cases where an autonomous vehicle has had an accident due to a failure to recognize a pedestrian, but there may be many causes like the result of errors in data acquisition, data fusion, or the decision-making algorithms themselves. The first pedestrian killed by an autonomous vehicle died because they were not recognized until too late and the car spent valuable time searching for an alternative route. The fact that the safety driver in the vehicle was also distracted and didn't react in a timely manner contributed. The conclusion in investigations was that software errors were contributing to the accident. Onboard computers make decisions based on the external and internal data; data fusion may or may not have been a contributing factor. Poor data frequently leads to poor decisions, but the various reports only discuss decisions, not the data that would be input to the decisions, fused or not.[6]

An additional source of data for autonomous vehicles can be a process known as floating car data where cars on a road can become active as moving sensors feeding information about traffic conditions to a traffic management center. This highlights a common data fusion problem. Coping with this amount of data will result in high CPU and network usage, but can build a more comprehensive model of the current environment the vehicle is acting in. CPU and networking loads will increase due the needs of analyzing structured data from on car telematics (easy) and unstructured data like camera feeds and data from street furniture (harder).

Cooperative systems such as collaborative robots should exchange data to improve safety and efficiency and fusing this data with other data will produce an improved model. The potential for damage and injury in a collaborative robotics scenario is high, when humans and robots are

[6] RAC. (2019, November 19). Uber self-driving car that killed pedestrian had software flaws. Retrieved March 27, 2020, from www.rac.co.uk/drive/news/motoring-news/uber-self-driving-car-that-killed-pedestrian-had-software-flaws/

working in the same space. Even with the limited intelligence robots in a warehouse, there is a high risk of problems. Autonomous collaborative robots increase this risk because there can be no fixed "safety area" if there is real collaboration between them and humans. Part of the risk mitigation would be for humans to have access to the same fused data that is modeling the world for collaborative robots, including sensor data from the robot, video, and LIDAR data as well as radar. The fusion of this data into a common model is vital for safety and effectiveness and would be shared between all collaborating participants.

Structured digital data is not the only type of data needed for a complete picture. Structured data is not sufficient or efficient enough to provide an accurate picture of the working environment or actions for the mobile robot or autonomous vehicle and will need different types and sources of data to make decisions. The sources of data are many and varied and some are mentioned in the autonomous vehicle example earlier in the chapter. Video data from cameras will be an important component that will add to the numeric structured data. Other sources of data include LIDAR and radar with similar goals to determine range, speed, and location of objects and GPS data for positioning. Unstructured data, such as video, audio, and analog, does not have a predefined data modal but can represent a large proportion of data that is collected and fused into the environmental model of a cobot.

Decisions using rules or AI based on structured data are relatively straightforward and can be automated readily as we discussed in Chapter 3, "Robotic Process Automation," on robotic process automation. Decisions that rely on more complex and unstructured data are themselves more complex, but they will need to develop decisions that will support the autonomy of collaborative robots.

Infrastructure Complexity Drives Data Fusion

If structured digital data is not sufficient for accurate decision-making in robotics and autonomous vehicles, additional data will need to be acquired to make the cobots safer or more reliable and decisions more accurate by improving the accuracy of the data set. To improve decision-making the data set needs nonstructured data that has to be gathered, processed, and fused with other data and used as part of the larger domain data set. Nonstructured data would improve performance and safety of a collaborative robot, for example, adding nonstructured data from a camera to data from radar will improve the model of the environment and ensure that the collaborative robot would better recognize obstructions. There is a real balancing act between effectiveness and complexity. Complexity in autonomous robots is increased by their evolution into collaborative robots since collaborative robots and human collaborators have to be in constant communication with each other and not reporting individually to a controller. To indicate the effect of collaboration on the complexity of data fusion, Figure 6-1 shows that robots who are collaborating with each other may be sharing updates to their internal representations of their environment as new data comes in and is fused into that internal representation. This will then generate updates to the collaborating human. The human may provide updates to the learned internal representation which has to be distributed between the robots and fused with their current representation. Data fusion is only one part of the picture, the robots and humans acquire data, and it is fused with other and existing data and used to create new internal mapping that is accurate in near real-time. A good example of this is a team of robots and a human building a wind turbine. Several robots are tasked with collecting and carrying turbine blades from a store and bringing them to the construction site, and as they bring a blade, it is handed to robots to hold the blade in position while the human attaches the blade to the turbine and makes any fine adjustments. All of the actors in the team, robot or human, needs to be in constant communication to enable each actor to make and modify decisions in a timely manner. Each of the actors is providing data that needs to be fused into the real-world model which can then be sent as an update to the team.

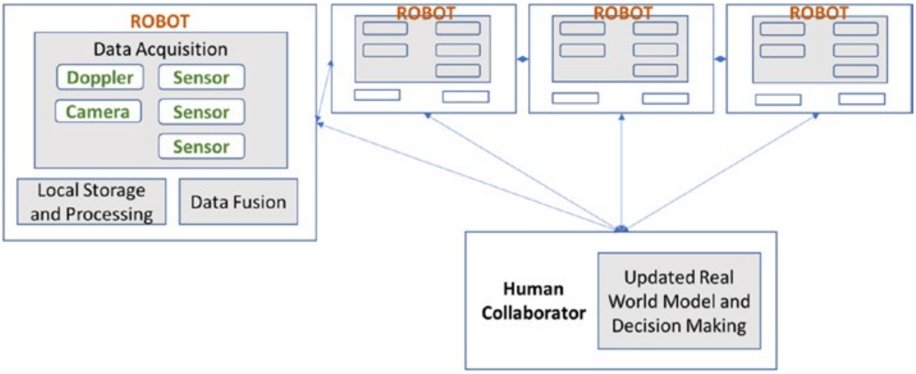

Figure 6-1. *Centralized data fusion architecture: collaboration and complexity*

Other modes of collaboration are noted and discussed in Chapter 4, "Robots in Teams," including the conflicts in a scenario where a single robot getting updates and instructions from multiple humans. Figure 6-1 will be used in this chapter to describe a centralized architecture for data fusion as well as to illustrate the complexity of the environment. Different data fusion architectures have an impact on performance, safety, and security in a collaborative robotics environment and will be discussed in the next section.

Architectures and Their Impact on Collaborative Robotics

Mobile robots can be considered as a set of sensors, cameras, radar detection, and so on. They are generating a huge amount of data as they move and complete tasks. The architecture of data capture and data fusion solutions can have an impact on the performance and reliability of the results that are sent to the decision-making part of a solution. There are two basic architectures for data fusion, a centralized architecture and a distributed architecture. Both bring advantages and disadvantages.

Figure 6-1 showed a simplified view of the complex environment that is prevalent in collaborative robotics. There are other representations of architectures for both data fusion as a process and robotics as an environment. In line with the focus of the book, we have only considered the use of multiple robots collaborating with each other. The box labeled ROBOT indicates the capabilities of a single robot that collaborates with other robots that may be different in physical appearance and construction but have the same capabilities. The smaller unlabeled boxes represent other robots identical to the structure used in the box labeled ROBOT and the text is omitted for clarity. The robot acquires data from sensors, cameras, and radar/LIDAR represented here by Doppler. As the data is acquired, it is stored in temporary storage and then uploaded to a central server where the data is managed and stored. The different types of data are fused together to develop an updated view that is a combination of the view that is common to both humans and robots, for example, a map of the warehouse layout, and new data that may indicate a change to the map, perhaps an obstruction. This updated view is then synchronized with the collaborating robots and humans so that a near real-time view can be used to improve activities. Any of the robots and humans can update what will become a new internal view of the environment. The robots still have autonomy even though they are relying on an identical common view.

Decision-making is not only part of each robot's capabilities but can also be centrally directed. A good example is task allocation while building a wind turbine. Each robot may be given different tasks such as carry blade or hold blade. The robot makes decisions on where to move or stand with no further reference to the central server unless it needs further instructions, for example, if a sensor detects an obstruction, the decision of the robot could be to stop and send new data to the server requesting alternate route instructions and informing other collaborators. It is clear that data volumes may be large in a cobotic world. Some sensors sample (read and transmit) data more than 1 million times a second. Having only ten sensors in a robot appendage would result in 10 million data points. Ten sensors may well be only the number of sensors in one part of a digit. Including unstructured data such as video and audio plus GPS data increases the amount of data, and this is only the tip of the iceberg. This, and all the additional data, can be processed or sent for processing to give the accurate picture of the operating environment for the robot. Fusing the large amount of data gathered will require high CPU usage to execute the data fusion algorithms, more CPU power than is available locally on the robot. Data transmission speeds and network resilience will be important in this scenario.

This scenario delivers accurate revisions to the internal mapping synchronized to each unit in the scenario, but this comes at a cost of stable high-speed networks and powerful servers. This is tenable in a controlled environment but may be more difficult to support in an unpredictable environment. There may also be the issue of low-powered devices that only have limited connectivity. Sensors at the end of a network will have to work with very low power. A sensor on a house light has access to plenty of power; a sensor that can only draw on power from the network or from intermittent solar power has to cope with low power. The sensors may not have the power available to carry out complex operations on data in this circumstance. Data may be summarized or analyzed on the spot before transmission.

Later in this chapter we will hear from a company that makes medical devices that have similar problems of low-power and unstable networks. These "edge of the network" devices have some capabilities that can overcome the high network and CPU utilizations of a centralized data fusion architecture. If you imagine a set of sensors in a home, some for measuring light levels, some for measuring temperature or humidity, and some for measuring movement, all of these send data back to a household controller. If this is the home of a vulnerable person, it is important that all the data is gathered and processed immediately and without the lag of updating a centralized server and waiting for analysis. Local household or even sensor processing is at the edge of the network and some analysis can provide immediate feedback. For example, if a motion sensor detects that the occupant has not moved for a set period of time instead of relaying the movement data, it can send an alert to the household or other supervisor that there has been no movement. This would be faster than sending data to a centralized supervisor for analysis.

A distributed data fusion architecture, as illustrated in Figure 6-2, conceptually needs to carry out all the storage and processing functions locally and then locally fuse the data into an updated real-world model. This model or the fused data to update the model can be sent to each actor in the collaboration team enabling them to use an updated model.

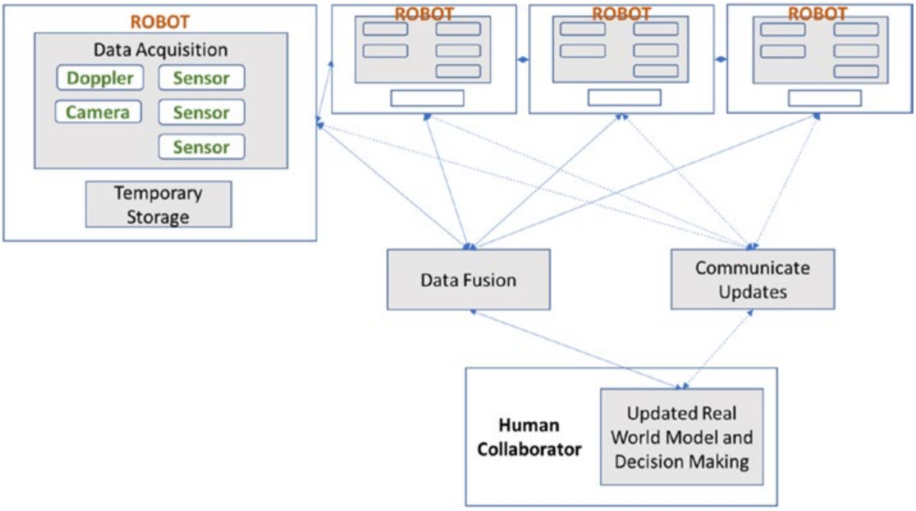

Figure 6-2. *Distributed data fusion architecture*

This synchronizing of communications between actors is a complex task. Newly updated models should not be able to overwrite data or models in other actors and should not have their models and fused data altered without checking if the data arriving is conflicting with the fused data or the updated local model. This is outside the scope of this chapter.

To reduce the network load and give near real-time results, each robot needs to be able to store data for partial processing and data fusion. The fused data can be used, along with any relevant external data, to update the internal mapping and decision base so that the results of the updates can be acted on locally. Updates can be shared, but using technology that can act locally or at the edge of the network and distributing the processing and data fusion means that each actor has a local accurate and timely view of the environment. For example, should a robot have an internal map of the environment, locations of shelves, height of shelves, width and length of pathways, and items to be collected, this will enable the robot to navigate through the warehouse. If there is an obstruction that is not mapped, this would cause problems, unless the robot uses sensors and cameras to recognize the obstruction. It can then update its own internal map. This obstruction then becomes a new part of the data set held by that robot. Local processing and data fusion enable this robot to react without delay and continue with a task even while it is sending the update to the rest of the team of collaborating robots. It does not need to send data for fusion; it has already accomplished this and just needs to pass the updates to the internal mapping.

Data Fusion Challenges for Collaborative Robots

Collaborative robots are autonomous and are designed to complete tasks in collaboration with other robots and human collaborators. They need to be aware of all of the actors, human or robot, that are working in the collaboration space. This challenges in the use of data fusion to maintain the real-world model of the working environment. We will address some of these challenges and use examples to illustrate the issues.

Problem Space

Once we move beyond the static, repetitive, manufacturing robots to autonomous mobile robots, the data challenges become evident. Static robots don't have to use sensors that map their relationship with the space they are working in. There are well-established safety procedures and physical barriers that can keep humans safe outside the space in which the robot acts. Sharing a space with a mobile robot requires both a human awareness of the robot's place in the environment and a robot's awareness of humans and other objects that share that space. In the case of warehouse robots, they may already have an internal, learned model of the environment, but unless they are designed to cope with an unstable environment, they may find it difficult to adapt to changes. These warehouse robots may only need a learned environment and a sensor that will allow them to sense when there are humans in the area. This does have some limitations as can be seen from the examples noted in the following text. Just moving around a learned environment does not require the sophistication of location awareness, sensing, data management, and decision-making that is the stock in trade of a collaborative robot.

Autonomous mobile robots have a requirement to adapt to an unstructured environment where objects, furniture, fixings, and people frequently move around the environment. These robots maintain a continual review of their location and its relationship to other objects in the real-time environment. In the case of warehouse robots, they may already have an internal, learned model of the environment, but they find it difficult to adapt to changes in their working environment without additional capabilities such as radar and cameras. The introduction of cameras as data acquisition tools creates additional demands on data fusion; structured data is not the only type of data in these environments. Increasingly nonstructured data has to be taken into account and used as part of the overall domain data set. This requires a different approach to managing and analysis of the complete data set.

Multisensor requirements and unstructured data such as audio or video need to be combined in a coherent way as key data in creating or updating learned and updated internal mapping accounts. Many additional types of data including audio, video, and multisensor will generate internal model updates describing the state of the robot. The large amount of data previously noted and the architectural challenges of low power and variable network connectivity combined with low storage capacity may become the norm in collaborative robots and have already been mentioned in our discussions about architecture. All of these challenges do not exist in a vacuum. One problem that needs to be addressed is that of lost, missing, or corrupt data. Bad data has the potential to cause dangerous or difficult situations. If a robot is moving across a warehouse floor and it has no information about the GPS location of other robots, it has to rely on its onboard sensors to help avoid a collision. If data wrongly indicates that the robot is handing an item over to another robot, this could cause safety issues and damage to the item, the robot, or the warehouse. This is already an issue in ordinary commercial computing but has a potential to introduce a dangerous element when autonomous mobile robots are working in collaboration and coexisting with humans in their environment. Missing or lost data may affect the accuracy of the data set on which an autonomous mobile robot will make decisions. In Chapter 7, "Robots in Society," we discuss military use of robotics and the debate that is current in the US military about giving robots the ability to take firing decisions.[7] The issues seen in different warehouse robots, medical robots, and military robots are similar even if the outcomes are not identical. Sensors that fail to transmit data or transmit unexpected data, too many outliers, and failed network connections all contribute to this problem domain. In any of these cases the problems could adversely impact performance of the robots or the collaborating team at the least, or cause a dangerous or costly situation to arise. There are the obvious problems with incomplete data yielding inaccurate analysis and possible poor decision-making. There is another area that you may have to consider and that is poor decisions based on data bias. There have been many articles regarding data bias in recent years covering a wide variety of cases ranging from sentencing of criminals to selection of CEOs.[8] In each of these cases the bias was found to be a result of the data samples selected. There are a number of cases where data selection can cause problems. A study of drug use in teenagers would generate false conclusions should the data set studied just use teenagers between 13 and 17. This would ignore use from the group of 18- to 19-year-olds who may have widely different usage since they are less likely to be in school.

[7]Fryer-Biggs, Z. (2019, September 3). Killer Robots and the New Era of Machine-Driven Warfare. Retrieved November 2, 2019.

[8]Lum, K., & Isaac, W. (2018). To predict and serve? Significance, 13(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Another area of study in the problem space is that of potential security holes. We have all seen movies where a hacker gains access to a robot and uses it for some sort of nefarious purpose. While there may be fiction, there are a number of reports of hacks to domestic robots and other Internet-linked equipment including one by James Vincent,[9] possibly due to generally low levels of security protocols and tools for this class of robots. The article described demonstrations of typical hacks in this area, where cameras were hacked to enable the hackers to spy on the householders. Perhaps more worrying were home robots controlled by outside sources and instructed to do some damage. Robots used for commercial projects have higher levels of security than domestic robots. The ENACT case study cited in the following text has as a major goal of securing healthcare critical robots and devices that use distributed and localized data processing, data management, and later data fusion.

Examples of Data Fusion Challenges

In the problem and architectural sections, we have discussed the internal mapping that a robot uses to understand its operating environment. We have also mentioned updates to that mapping and the subsequent state changes involved. In this section there are two examples that demonstrate the future need for robots that are more aware of the changing state of their environment and the robot's ability to create updates to internal mapping. Both of these examples are widely reported and our use of them is only to illustrate the changes needed before autonomous mobile robots can operate safely in an unknown and possibly changing operating environment. Cobots with additional detection capabilities can use them to generate a more accurate internal data model of the environment allowing decision-making to resolve the problem before it occurs.

[9]The Verge, James Vincent: Accessed in November 2019, published August 2017 www.theverge.com/2017/8/22/16183514/hack-home-robot-surveillance-ioactive

Example 1

A household robot vacuum cleaner will traverse house floors picking up debris. This robot uses infrared to detect walls and obstructions and this will help it navigate around the room. If the robot detects a new obstruction, it will try to navigate around that new obstruction. If the cleaning is set to begin at a time when there are no people, cats, or dogs in the room, then this works well. The problem comes in when nondetectable deposits are made that the robot can't recognize. In one case[10] the owner's dog had an accident, leaving feces on the floor. The robot started to clean, rolled over the faces, and carried on cleaning with the wheels dirty. This spread the deposit around the room and probably ruined the owner's day. The use of video and image recognition could be added to the robot allowing the fusion of data to help in the decision-making process by updating the internal model of the robot.

Example 2

A similar accident happened in an Amazon warehouse.[11] In this case an automated robot punctured a canister of bear repellant that fell off a shelf according to news reports. Employees ended in hospital as a result of this incident. If the robot had been able to recognize the canister, perhaps that it contained hazardous material, it could have stopped or avoided the obstruction and updated other robots to enable them to do likewise. The capacity to recognize a problem from data gathered, perhaps by video, and fusing this data into the data model would have created an internal model that could recognize the hazard and the changed risk profile would then generate a decision to avoid or evade the hazardous material. Again,

[10] www.theguardian.com/technology/2016/aug/15/roomba-robot-vacuum-poopocalypse-facebook-post

[11] www.theguardian.com/technology/2018/dec/06/24-us-amazon-workers-hospitalised-after-robot-sets-off-bear-repellent

this indicates that awareness of surroundings and changed state of the area of robot operation could have led to a real-time update to the internal representations of the warehouse which would have the potential to avert a nasty accident.

Potential Solutions

Problems fall into two broad categories, coping with an unstable environment and coping with data. This is true for the two examples as well as for collaborative robotics. Before looking at the technical solutions, we can consider operational solutions to the preceding two examples. In both cases there are operational solutions that can be implemented immediately. In the case of the household robot vacuum cleaner, the owners can restrict the operation of the cleaner to rooms or times where the likelihood of a pet generated, or other accident, is not possible. The possible solution for dangerous substances is to put these substances in a different facility and introduce changes to the storage, handling, and management processes. Both of these may only be "band aid" are likely to be faster and more cost effective in the short term than retooling or purposing of automated robots.

While there are operational solutions giving short-term fixes, the short-term changes would not be sustainable in the longer term. It is reasonable to expect that future robot vacuum cleaners would have the ability to recognize that they are making more mess than they remove. Robots in a warehouse should be able to carry a map of what robots and humans there are and identify the goods being carried by either. The long-term solution would involve the development of collaborative robots in both scenarios, able to move independently, and recognize other actors in their space. In both these examples, the robot should be able to take decisions to mitigate the risk of continuing their task by considering their internal map; again this would require accurate and extensive data gathering and data fusion. While these more advanced solutions involve cost, creating newer more autonomous robots would return the business model flexibility. Returning flexibility to the business model by having all the warehouses equipped to the same standard instead of having different warehouses for different good would reduce costs in the long term and enable changes to the business model to be made faster and easier. Ensuring that operation of a robot vacuum cleaner could be as simple as charge it up and let it go anywhere is likely to make these more attractive to buyers and improve sales.

Focusing on collaborative robots, the problems of creating and maintaining a shared awareness of environment, tools, and tasks by using data from sensors and cameras require the use of data management and fusion. This will provide both knowledge and context for a collaborative robot. In Figure 6-3 the completed process has been described in numbered boxes. The numbers in this paragraph refer to the box number in the figure. The process can start (box 1) with raw data acquisition where signals are converted into machine readable form by the sensors. There are few standards in this area, and this will generate additional work for international standards bodies determining agreed communication protocols and agreed structures and capabilities for the next stage of data acquisition (box 2) where data is stored or transmitted depending on the architecture that has been chosen (box 3). Data processing (box 4) follows data acquisition and depends on numerous factors such as the type of sensor, the capabilities of the device that carries the sensor, and the types of data being collected and environmental conditions. Once the raw data is stored, it can be examined and combined with other data by data fusion (box 6) with other data to create information. This information can then be fused with other data (box 6) to create a more comprehensive data set to provide information for analysis and decision-making (box 7). As we have mentioned there are storage costs and processing costs that need to be reconciled alongside networking costs. When designing a collaborative robot architecture, all of these factors will influence decisions on sensors and architectures. Should the architecture decision be to use mobile device communications styles, the challenges of storage and processing can be addressed by a number of techniques, although there may be changes to this strategy going forward with the development of 5G networks. 5G networks can connect different devices at high speed; it is claimed that 5G networks are 100 times faster than 4G networks, although this is at the theoretical maximum performance. This would mean that more data can be transmitted faster and this may remove the need for segmentation. However, at the current time the rollout of 5G technology is slow and patchy, so there is a need for managing large amounts of data in the current networking environment. If there is a large amount of data, it is possible to use segmentation techniques to create smaller coherent sets of data that are more suitable for the processing power of the devices. To resolve the storage issues, data for context information can be processed. With the context information stored, the raw data can be discarded and saving space and reducing networking traffic.

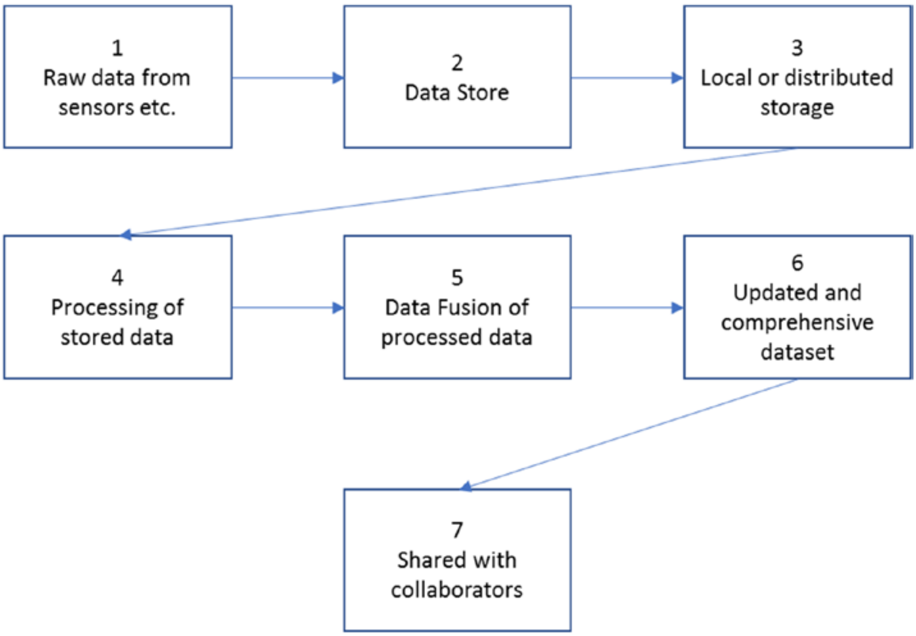

Figure 6-3. *Process flow for maintaining a shared environment*

Data processing, as the name implies, involves subjecting raw data to some form of handling that give a level of information. Taking a set of temperature readings and converting them into classifications such as "too hot or too cold" is typical of the level of data processing that may be done at the edge of the network. Other processing may be comparing the location of the collaborative robot to the mapped safe area to give a simple evaluation of safe or not safe. Again, this type of context information can be either used to generate an action; in the case of the robot not being in a safe area, the robot may decide to stop moving until it is given fresh instructions. This can either happen at the edge of the network or this context data can be transmitted to a central server. In the case of safety, it is more likely that the processing will be local; in the case of ambient temperature, the data may be centrally processed and matched with other data from other robots enabling the central location to take a decision on altering the ambient temperature.

The introduction of cameras or other video or audio equipment into the monitoring mix can add to the processing burden. Unstructured data analysis is more complex than the structured data analysis referred to earlier. Each data set that is processed needs to be related to a view of the whole environment and this can only be achieved by some level of data fusion. Fusing video, radar, and sensor data can give a more comprehensive view of the world in which the robots and humans act in. Fusion of video data with structured data is done by using algorithms that combine, reconcile, and relate different types of data to create a complete context. The data might show, for example, that one robot is in the shade and other robots are in sunshine. This may be the only reason for the differences in temperature sensed by the robots. If the robots in the sunshine are approaching the limits of operational safety, an imperative command could be transmitted to the robot team to all move to the shade to stay cool enough for operational safety.

While these different types of data and processing can yield important safety or task results, we also need to consider other problems with data. The repair of missing or incorrect data including outliers is a data science problem but we consider it here because of the potential for safety violations. There are many algorithms for identifying outliers in structured data and these can be applied to data as it is gathered. An outlier, or a value that is outside the normal range of the data values in a data set, can skew machine-learning results. Some machine-learning algorithms are sensitive to the distribution of data. There are a wide variety of preprocessing algorithms to detect outliers and repair data sets such as extreme value analysis. Other problems with data such as missing data, perhaps caused by a camera failure, can be repaired or updated by deriving missing images using data from other cameras and data such as GPS location.

Impact on Collaborative Robotics

Autonomous vehicles are solving some of the problems that we mention here; however, the solutions are more focused on safety and transport rather than interacting and collaborating with humans. Data fusion is more necessary in collaborative robotics than in autonomous vehicles since human collaborators need a “world” view that is synchronized and supplies the same data that the robots interchange. Missing, inaccurate, or poor data can influence the decisions that the humans and robots take.

Incomplete or false data fusion can have serious consequences and more research into data acquisition, data processing, and decision-making will help overcome this. Security and privacy are also factors to be considered. In the world of working robots, security has to prevent unauthorized access, or changes to data that can influence decisions have to be prevented or detected and mitigated. If robots have cameras to assist in their work, the camera may pick up other images unintentionally. There have been amusing images of photo bombers and the more sinister lack of control in robots moving around the home with a camera capturing and transmitting image data of private moments. In some domains, such as healthcare or medical applications, there is also the consideration of privacy. This is discussed more fully in the ENACT project research notes later in the chapter.

Supporting Research Projects

In this section we explore two research projects who have been investigating data fusion and related topics from their own perspective, having an impact on collaborative robotics and ultimately the future of work.

ENACT

ENACT is a research project that is handling data processing at the edge of the network bringing some perspectives that are important for collaborative robotics. ENACT[12] is a project funded by the European Union under the H2020 Programme[13] and is expected to end in January 2020. Although the project is focused on the Internet of Things (IoT) technologies in transport and eHealth, its significance in this chapter is the development of research into edge of the network technology. A robot that is engaged in search-and-rescue operations in a hostile environment may be considered a device at the edge of the network and may be using similar technologies to mobile, low-power devices. Low power or power conservation would be important to a robot that is away from power supplies and needs motive power to complete its tasks.

[12]Consortium, E. (April 1, 2020). ENACT: Development, Operation, and Quality Assurance of Trustworthy. Retrieved April 1, 2020, from www.enact-project.eu

[13]Commission, E. U. (April 1, 2020). What is Horizon 2020? | Horizon 2020. Retrieved April 1, 2020, from https://ec.europa.eu/programmes/horizon2020/what-horizon-2020

The ENACT project will develop, as one of many use cases, a digital health system for supporting and enabling patients to remain in their home instead of entering a care facility during either treatment of care. In this section we will consider elderly care as part of this use case. This use case controls lifestyle equipment, light and heating controls, door locks, and so on. the use case also controls various types of medical devices and sensors including blood pressure, weight, fall detection, and video surveillance equipment and sensors. This system needs to integrate with other systems, like management and emergency systems, and provide information/alarms to response centers and care givers. All of this is done at the edge of the system using a variety of networking tools. Normal computer networks will have a computer at the end of the network; however, the sensors in this use case have very little processing, storage, and networking power. Many of these devices are single-function tools that measure temperature and so on, and their data needs interpreting by other equipment, either a local or distributed computer. This type of architecture is not only increasing in use but is related strongly to the use of robots that have a tenuous link to their networks. In these scenarios the data processing may include data fusion as a component, although in the early stages of ENACT, data fusion is a future goal.

EXPERT INTERVIEW WITH ARNOR SOLBERG

Dr. Arnor Solberg from Tellu is a major contributor to the ENACT project. In an interview with him, we discussed his edge of the network technology prototypes, his view of the relevance of data fusion, and what his plans are for the future of the ENACT and other projects.

- There are several issues that elderly encounter when living in their own home that can be managed and monitored and have until recently required a part-time or full-time caregiver. The same issues are present even if they move into a nursing home, for example:

- **Falling**
- **Being too hot or too cold**
- **Failing to eat**
- **Blood pressure problems**

- With low-cost low-power devices, much of the in-home monitoring can be done by sensors, video cameras, and an application that gathers and analyzes data. The analysis is carried out centrally.
- Some of the analysis can be complex, asking questions like “Is the patient lying on the sofa because they are tired or because they are ill or faint?” This can be resolved using a combination of tools, measuring blood pressure, checking the temperature of the room, last food consumption times, and so on. It may be necessary to contact the patient by mobile or some other communication device. This leads to the concern that there are too many false alarms which absorb resources and annoy the patient. There is also the possibility of missing or false data. A failed sensor for room temperature would be critical if it was the only sensor in the room. In some houses the location of the heating thermostat can decide the temperature of only one part of the dwelling and opening or closing doors can have a significant effect on temperature.
- If the patient is in the main body of the dwelling and can be seen, much additional data can be gleaned from video or other monitoring. If the patient is hidden, for example, behind a sofa sensor, monitored data will need to be analyzed further to establish if the patient has collapsed or is picking up something that has dropped behind the sofa. Again, this can lead to false alarms.

- Much of the important data comes from sensors and monitors deployed around the house and uses device management and where possible remote deployment of software. The use of processing at the edge of the network will lead to better security and privacy, with only the context relevant data being sent to a central location for analysis. The current application uses rules to decide on alerts and the rules are use case dependent.
- Currently a person monitoring the patient uses alarms as a guide to action; however, there is no integrated view of the data. In the future Tellu will be developing a data fusion approach that will enable better accuracy and the ability to gain a holistic view of the data in a case. This should remove more positives and improve overall performance. At the moment it is not decided if the data fusion algorithms will be deployed locally on the devices or centrally prior to analysis.

CVDI Co-Botics Project

The research project Co-Botics[14]—intelligent cooperating robots and humans: parts 1 and 2—has been managed by Professor Moncef Gabbouj of Tampere Technical University in Finland. The project was initiated as part of the Center for Visual and Decision Informatics (CVDI) is a US government National Science Foundation (NSF) initiative, and a number of commercial and academic partners collaborate, generating funding and research partnerships. The initial goal of the project was to investigate

[14]Gabbouj, M. (2020, August). Co-Botics—Intelligent Cooperating Robots and Humans—Phase II 7a.028.TUT—CVDI. Retrieved April 1, 2020, from www.nsfcvdi.org/wordpress/cvdi_project/co-botics-intelligent-cooperating-robots-and-humans-phase-ii-7a-028-tut/

computational models to automate and monitor interactions between humans and robots. The models would provide a view of reactions to various behaviors. This is explored more fully in Chapter 7, "Robots in Society." The project was focused on collaborative robotics and used the coined term Co-Botics. The Tampere group has focused on advanced machine learning and pattern recognition to facilitate intelligent shared cooperation between robots and humans. The most significant for this chapter is the research into multiview data analysis that can describe cues from the real world. Ultimately the project intends to combine visual information analysis with sensor data analysis and use this combined analysis for decision-making.

The project will continue toward enhancing the performance of multimodal visual/sensor data analysis methods for efficient robot-human interaction in efficient scheduling applications. Moreover, it will focus on creating data visualizations that combine information coming from various types of sources (visual, depth, audio) in order to provide insights on the way robots perceive their environment. We believe that such visualizations will allow us a better understanding of how to enhance the overall operation and increase intelligence of robotic units in the targeted scenarios.

EXPERT INTERVIEW WITH MONCEF GABBOUJ

We asked Professor Gabbouj a number of questions, via email, to explore the topics and we report on his answers here.

- Deep learning has been defined in Chapter 2, "Technology Definitions," and is used here in the context of the number of learning layers that transform raw data into higher abstractions. The synergy of the topics of multiview learning and learning to rank can be used in multiview lingual text ranking and image data ranking. Image data ranking will be beneficial in data fusion applications in collaborative robotics. Experimental results are promising and show that there is a performance differential between supervised and unsupervised models, although this is still theoretical work. There are a number of papers that support and describe this work[15,16,17] and are recommended reading for more information on the topic.

- In discussion about other components in a robotics solution, Professor Gabbouj noted that there are two additional components in robotics solutions beyond data fusion. They are co-creation and situation awareness. Co-creation refers to the human-machine close or intertwined collaboration. Situation awareness is about developing the actual AI solutions for this awareness.
- A discussion about ineffective fusion of structured and unstructured data led to a comment that poor data fusion cannot only produce bad results but may also lead to wrong conclusions if the results highlight any biases often present in training data (biased training data is discussed earlier in the chapter). Proper data fusion must then be verified in varying environments to make sure the solutions are robust.

[15]Guanqun Cao, Alexandros Iosifidis, Moncef Gabbouj, Vijay Raghavan, Raju Gottumukkala, Deep Multi-view Learning to Rank, IEEE Trans. on Knowledge and Data Engineering, 20 Sept. 2019. arXiv:1801.10402.

[16]Guanqun Cao, Alexandros Iosifidis, Moncef Gabbouj, "Multi-modal subspace learning with dropout regularization for cross-modal recognition and retrieval," 6th International Conference on Image Processing Theory, Tools and Applications, IPTA 2016, 12-15 Dec. 2016, Oulu, Finland [Winner of the Student Best Paper Award at IPTA 2016].

[17]Guanqun Cao, Iftikhar Ahmad, Honglei Zhang, Weiyi Xie and Moncef Gabbouj, "BALANCE LEARNING TO RANK IN BIG DATA," 22nd European Signal Processing Conference, EUSIPCO 2014, 1-5 September 2014, Lisbon, Portugal, pp. 1422-1426.

- Tainted or biased data or occasions when different sources of data contradict or are inconsistent with each other are common data science discussions. Putting these questions in a robotic context, Professor Gabbouj noted that it is not necessarily due to bad or tainted data; it could arise from different sensors measuring different characteristics of a system or those different characteristics may have conflicting implications and he mentions that some perceptual illusions in humans arise in this way. In the era of big data, a lot of errors come into play: wrong data, missing data, biased data. The good news is that we have ways to deal with some of these issues. For modalities to infer from available data (and not about the missing data), we can also prevent data overfitting at training via mathematical regularization, for example, our work in the paper mentioned in Chapter 2, "Technology Definitions."
- Professor Gabbouj mentioned another issue about balanced learning in large data and their proposed distributed learning to rank method. It is not applicable to train a centralized ranking and the distributed methodology can easily be scaled up to billions of images. Experimentally the proposed method outperforms a straightforward aggregation of boosting algorithms.
- Returning to the topic of missing or inaccurate data, Professor Gabbouj described a situation where we know that in some applications some data is missing and, in some cases, it can be inferred from available data. In other occasions, we simply ignore the missing data and do the inference based on what is available. Data filtering seeking outlier removal is often applied, especially if we know ahead of time that such outliers do not arise in normal situations.

- Other critical problems are many, for example, what data and what modalities are sufficient to perform inference/fusion/learning, what are the best models, and in what sense are those models best.

These points in the interview are interesting in the context of collaborative robotics, even though they only address a few of the software issues in this space. The point that Professor Gabbouj made regarding human/robot fusion (co-creation and situational awareness still to be addressed) indicates that in this area collaborative robotics will need much more research before the collaborative robots mentioned in many scenarios become a reality.

Summary and Conclusion

Looking into the first technology challenge, we explored the use of data fusion, an enabling technology that can help solve one of the hardest problems in collaborative robotics. Collaborative robots need a shared, common view of the unstable environment that they inhabit and this view would not be possible without data fusion. After defining data fusion, we looked at the leading example of data fusion, the human being. Humans are highly successful data fusion experts. They can pull together data from many inputs, sight, sound, touch, and all the other senses to build a picture of their surroundings and enabling them to relate to those surroundings. In data fusion terms, collaborative robots should have the human data fusion ability although without converting; indeed, data fusion success should, in the future, be measured against human capabilities. Another area where humans excel and few robots can come near is in pattern matching. Humans can use very sparse information to match patterns, such as a limited side view of a person's face. In pattern matching humans are adept at identifying patterns when large amounts of data are missing.

In robotics structured digital data is reviewed as a relatively easy way of describing a subject's state. This is not the only type of data to be considered. Structured digital data needs to be added to video, audio GPS, and other data to create an almost complete version of the autonomous mobile robot's operating environment. These are the bigger issues in data fusion that arise for a need for more comprehensive internal maps for autonomous mobile robots. The robot can also be considered a mobile "edge of the network" device that may have limited power and networking capabilities and needing special treatment of data to improve data quality. Data fusion can help overcome some existing robot's limited adaptability. If the robots that we mentioned in the examples earlier in the chapter had more sensors, cameras, and other data-gathering equipment, they could respond effectively to their environment rather than merely moving around without the context of the environment to help them overcome obstructions. A mutually understood internal map or environment is not just the goal for the robots; their human collaborators also need to understand the environmental context of all of the collaborators, both human and robot. This mutual understanding should enable the collaborating team to operate in safety and with effective communication to complete their mutual tasks.

Research is important in moving data fusion forward and two research projects were identified as relevant. The relevance of one of the projects, ENACT, was in its forward-looking resolve to solve problems identified in their current prototypes. Their prototypes are already on trial in healthcare sites and the data collected is rudimentary at the moment. But the plan is to upgrade the analysis of existing and future data by data fusion of sensor and other data to improve the accuracy of the current system. This upgrade would also improve the quality of decision-making and improving the quality of life for the elderly residents, allowing them to stay in their own homes longer.

The second research project, Co-Botics, in partnership with the Center for Visual and Decision Informatics is researching ways of both fusing and analyzing data faster and more accurately than is currently possible. This research pioneered new ways of analyzing data that starts with evaluating a single data type analysis and then introducing a second data type and a third fused together and analyzed. In doing this the ability to analyze new data sets from multisensors will improve. Once the research moves into the applied phase, practical applications will come. This research is ongoing but is not concluded. Building these tools into a working robot environment may be a long-term goal. Solving data fusion problems will be a major milestone in collaborative robotics. Other areas of research and development are underway; using data fusion in the intelligent transport world is underway, focused mostly on autonomous vehicles. Vehicle manufacturers are working hard to solve the fusion of data from vehicles and roadside sensors. Chapter 5, "Robots Without Arms," expands on this use case.

Data fusion is an important component of collaborative robotics and it is not expected to be fully usable in the short term, but research is improving the prospects of having this solution available in the midterm. Although it is an important component and essential for collaborative robotics to work, there are other problems to be solved as well and these are discussed in future chapters. Using the human as an example of good data fusion shows just how far away an ideal solution is. Humans and robot common world views are still a long way off and as such the level of autonomy in collaborating robots is still set at a low level.

CHAPTER 7

Robots in Society

Corporate Responsibilities

> *Unlike traditional injuries, tort law will have difficulty finding the injuries caused by highly sophisticated AI to be the fault of someone's negligence or some product's defect.*[1]
>
> —Yoshikawa, J. (2018)

Consider the humble vacuum robot. In Chapter 1, "Will Robots Replace You?", we recounted the story in which a friend's Roomba navigated itself into and through a cat litter box and across her wood floor, depositing and spreading litter and stool. In her case, the damage was temporary, but what if the damage had been permanent, perhaps destroying an antique rug? If we assume that the robot is a tool, we might ask whether the scenario was foreseeable and whether the manufacturer was strictly liable because product warnings and instructions failed to clearly warn the user about floor-level containers with a low barrier (in this case a litter box).

If we think of the robot as an autonomous agent of a service provider, or if the robot were remote controlled by a human, would our legal response change? Robots such as autonomous vehicles and surgical or

[1]Sharing the Costs of Artificial Intelligence: Universal No-Fault Social Insurance for Personal Injuries. *Vanderbilt Journal of Entertainment & Technology Law,* 2018; 21:1155.

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_7

search-and-rescue cobots will occasionally commit errors that directly or indirectly lead to injury or even death. The fact that they might make fewer errors than humans in the same situation may justify their use, but it will not free them from liability.

In this chapter we will discuss the policies and tools for regulating, monitoring, and governing collaboration between automated processes, robots, and humans. We will discuss these work-related issues in the following sections:

- *"*What Can Go Wrong?*"*, which examines some the incidents involving robots and automation
- *"*Legal Remedies,*"* which considers some of the legal and policy issues surrounding robotic activities
- *"*Robots in Corporations, Corporations in Robots,*"* which proposes guidelines for businesses and societies that create and use robots

What Can Go Wrong?

The example of the vacuum robot is trivial when compared to the tragedies that can occur when automation or robotic systems fail. A fatal accident involving Tesla Model S's autopilot system occurred in 2016. According to an IEEE article on the accident, the cameras and radar used by the autopilot system were not able to recognize, "the white side of the tractor trailer against a brightly lit sky," and according to the CEO, "[the] radar tunes out what looks like an overhead road sign to avoid false braking events."[2] In their statement

[2]Cited from E. Ackerman (July 1, 2016). Driving Car Crash Reminds Us That Robots Aren't Perfect, *IEEE Spectrum*. https://spectrum.ieee.org/cars-that-think/transportation/self-driving/fatal-tesla-autopilot-crash-reminds-us-that-robots-arent-perfect [accessed on March 8, 2020].

concerning the accident, Tesla noted that accidents involving their automobiles are extremely rare, and asserted that:

> *Every time that Autopilot is engaged, the car reminds the driver to "Always keep your hands on the wheel. Be prepared to take over at any time." The system also makes frequent checks to ensure that the driver's hands remain on the wheel and provides visual and audible alerts if hands-on is not detected. It then gradually slows down the car until hands-on is detected again.*[3]

The earliest recorded instances of deaths due to robot failures occurred in the late 1970s and early 1980s. In 1979, Robert Williams, a worker at a Ford Motor casting plant, was asked to climb up to a shelf to retrieve castings because a five-story industrial robot which moved castings on and off the rack had provided workers with possibly erroneous information on the number of parts.[4] While on the rack, a one-ton cart with a mechanical arm hit and killed Williams.

In 1981, Kenji Urada, a Kawasaki employee, was killed when he entered a restricted safety zone to perform maintenance on a robot that he had failed to completely turn off. The robot's hydraulic arm pushed him into adjacent machinery. "According to factory officials, a wire mesh fence around the robot would have shut off the unit's power supply when unhooked. But instead of opening it, Urada had apparently jumped over the fence."[5]

---

[3]The Tesla Team (June 30, 2016). `www.tesla.com/blog/tragic-loss` [accessed on March 8, 2020].

[4]`www.nytimes.com/1983/08/11/us/around-the-nation-jury-awards-10-million-in-killing-by-robot.html` [accessed on March 8, 2020]. Latter articles about the event claim that the robotic system was acting too slowly, and Mr. Williams "… was reported to have climbed into the storage rack to retrieve parts manually" (`www.forbes.com/sites/theopriestley/2015/07/02/is-this-a-killer-robot-uprising-hardly/#2ebc7f7d2396` [accessed on March 8, 2020]).

[5]From the archive, 9 December 1981: Robot kills factory worker. `www.theguardian.com/theguardian/2014/dec/09/robot-kills-factory-worker` [accessed on March 8, 2020].

More recently, in 2015 Wanda Holbrook, a maintenance technician for Ventra Ionia Main, was killed while performing routine activities when she was trapped and crushed to death by robotic machinery. The lawsuit filed by her husband named five robotics companies, FANUC, Nachi, and Lincoln that made the robots and Flex-N-Gate and Prodomax that helped with installation and servicing. The lawsuit claimed that the company's safety systems had failed: the robot should not have been able to enter that area of the factory, and the "safety doors that were installed specifically to prevent robot movement were not effective."[6]

Several patterns emerge from these disasters: (1) data from sensors and databases can contain errors that lead to fatal decisions by humans and machines; (2) except in the case of Tesla Model S, the machines had little ability to detect and self-regulate their behavior when humans were nearby; and (3) humans need to be vigilant when robots are nearby or in charge, and workers need training on safe and appropriate procedures (including how to shut down the robot). Unfortunately, as of 2020, the Occupational Health and Safety Administration (OSHA) "currently [has] no specific OSHA standards for the robotics industry."[7]

In addition to physical injuries, other forms of injury may occur. A robot might incorrectly classify an employee's behavior as dangerous, video record the instance, and send the embarrassing video to the

[6]Courthouse News Service (March 8, 2017), www.courthousenews.com/work-robot-blamed-michigan-womans-death/ [accessed on March 8, 2020]. See also E. Livni (March 13, 2017). A rogue robot is blamed for a human colleague's gruesome death, *Quartz,* https://qz.com/931304/a-robot-is-blamed-in-death-of-a-maintenance-technician-at-ventra-ionia-main-in-michigan/ [accessed on March 8, 2020].

[7]Occupational Health and Safety Administration (accessed on March 8, 2020), Robotics, www.osha.gov/SLTC/robotics/standards.html. However, the *American National Standards Institute (ANSI) and the International Organization for Standardization (ISO) have been developing guidelines*, for example, https://webstore.ansi.org/preview-pages/RIA/preview_RIA+TR+R15.606-2016.pdf, and Collaborative Robot Safety, www.iso.org/standard/62996.html.

employee's manager. Depending on the circumstances, this might be a privacy violation, and even if it is not considered illegal, it would certainly raise issues of trust. These possibilities suggest the need for government statutes and regulations and corporate policies that go beyond typical safety standards.

Can Robots Be Ethical and Self-Regulating?

Ethics can be defined as "the study of what is morally right and wrong, or a set of beliefs about what is morally right and wrong."[8] To act ethically, individuals or organizations must monitor and regulate their own behavior.

One approach to regulating the behavior of robots proposes that robots should be programmed with a set of ethical principles that govern their decision-making, or that they acquire ethics through experience and human coaching. The most famous rule-based approach is from Isaac Asimov's *Three Laws of Robotics*, first described in 1942 short story, *Runaround*:

> *First law*: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
>
> *Second law*: A robot must obey the orders given to it by human beings, except where such orders would conflict with the first law.
>
> *Third law*: A robot must protect its own existence as long as such protection does not conflict with the first or second law.

[8] Definition of ethics from the Cambridge Academic Content Dictionary, Cambridge University Press. https://dictionary.cambridge.org/us/dictionary/english/ethics [accessed on March 30, 2020].

As Asimov explored through his fiction, the flaws and contradictions that arose from the application of these laws, a fourth, or 0th law, were added:

> *Zeroth law*: A robot may not harm humanity or, by inaction, allow humanity to come to harm.

These laws are the source of many fictional accounts of anthropomorphic robots, but we are not aware of any attempt to embed these specific laws into working robotic systems. Indeed, these laws were designed to fail in interesting ways, enriching Asimov's brilliant storytelling. They are inherently adversarial, vague, and based on flawed theories of ethics[9]: adversarial because they implied that without these laws, robots might intentionally harm or dominate humans; vague because the term *harm* is inherently vague—how much risk is considered harmful and how are harms and benefits balanced? Humans frequently engage in activities that are beneficial but involve risk, for example, playing competitive sports or caring for ill.

Inspired by Asimov, others have tried to propose models of ethics that could be implemented in a machine.[10] These approaches suffer from two basic flaws: (1) categorizing behavior as ethical or unethical can be difficult—is it sometimes okay for a collaborative robot to lie to a medical patient or team member for their own good?—and (2) rules for complex systems are typically incomplete and are difficult to maintain or evolve when confronted with unexpected situations.

Regarding the first flaw, for millennia philosophers have been attempting to systematize ethics into a single unambiguous philosophical model, it is unlikely we will achieve a breakthrough in the next fifty. Deep

[9]Dvorsky, G. (2014). Why Asimov's Three Laws Of Robotics Can't Protect Us. *Gizmodo*.

[10]Arkoudas, K., Bringsjord, S., & Bello, P. (2005, November). Toward ethical robots via mechanized deontic logic. In AAAI fall symposium on machine ethics (pp. 17-23). Menlo Park, CA: The AAAI Press.

wisdom has been achieved in religious and secular explorations, but there is no single system that everyone agrees to, across all identity groups and for all contexts. Specifying the rules for ethical behavior for a machine that is physically (and perhaps logically) more powerful than humans and that obeys the rules literally and at scale (millions of robots following the same logic) is a frightening challenge.[11]

The second flaw, which addresses the inherent limits of the rule-based systems, applies not only to ethics but to many situations that require complex social interactions and that involve unexpected variations. We saw this in our earlier discussions of RPA, chatbots, and collaborative robots. Even in relatively simple business process applications, such as payroll or database security, where the rule-based systems might contain thousands of rules and exceptions to those rules, the sequence in which the rules are applied is critical—inserting a new rule or deleting an old rule can have unintended consequences. It is unlikely that any set of rules, which are to be literally and unambiguously interpreted, can cover every circumstance or consider every side effect.[12]

If rule-based systems are not enough for ensuring ethical and appropriate behavior, what about deep-learning systems? AlphaZero, a deep-learning system, has taught itself to play two very different board games, Go and chess, at superhuman levels.[13] Can deep-learning systems acquire ethics on par or even superior to humans?

---

[11]See Muehlhauser and Helm's delightful discussion of the Golem Genie which is super-powerful and literal-minded, in Muehlhauser, L., & Helm, L. (2012). The singularity and machine ethics. In Singularity Hypotheses (pp. 101-126). *Springer*, Berlin, Heidelberg.

[12]Yampolskiy, R. V. (2013). Attempts to attribute moral agency to intelligent machines are misguided. In *Proceedings of Annual Meeting of the International Association for Computing and Philosophy, University of Maryland at College Park, MD*.

[13]The Straits Times (March 15, 2016). "Google's AlphaGo gets 'divine' Go ranking." www.straitstimes.com/asia/east-asia/googles-alphago-gets-divine-go-ranking [accessed on March 8, 2020].

Unlike current AI systems, humans learn ethics and good behavior through experience and from parent, peers, and teachers. Humans are not general-purpose computers, and our cognitive ability in certain domains (such as linguistic communication) is equal or superior to Turing machines.[14] Human evolution occurred with specific memory, sensory, and attentional limits and with specialized neurophysiology for surviving in a dangerous world. Our brains and intelligence are specialized for social interaction and survival and, at the same time, are capable of general-purpose analysis, generalized learning, and empathy. How this is achieved is still a mystery.[15]

For current deep-learning systems to learn, they must be provided with an *objective function* that specifies how to recognize correct or incorrect outcomes (such as winning a game of *Go*) so that their behaviors can be positively or negatively reinforced. Defining moral behavior as an objective function is an interesting challenge for machine-leaning researchers, but as we argued earlier, philosophers cannot agree about those principles. There are differences in ethics across cultures, and there are many edge cases that are difficult for humans in the same culture to agree upon.

Moreover, current machine-learning approaches will not help prevent the scenario described in the beginning of the chapter, involving a vacuum robot. Nor will it address the issues surrounding liability or overcome

[14]Chomsky's analysis of the cognitive requirements needed for language suggests that Turing machines, "the pinnacle of all possible mathematical machines … is also the minimum needed for human cognition." Waldrop, M. M. (2001). *The Dream Machine: JCR Licklider and the Revolution That Made Computing Personal*. Viking Penguin. p132.

[15]Thomas, J. I. (2019). Current Status of Consciousness Research from the Neuroscience Perspective. *Acta Scientific Neurology*, *2*, 38-44.; Grossberg, S. (2019). The resonant brain: How attentive conscious seeing regulates action sequences that interact with attentive cognitive learning, recognition, and prediction. *Attention, Perception, & Psychophysics*, *81*(7), 2237-2264.

the fundamental flaws—the vacuum robot lacked the sensory abilities to detect the effect of its actions and the reasoning abilities to infer causal connections between its actions and environmental changes. Humans and other animals evolved to survive in a natural and largely uncontrollable environment and, through evolution, have gained the sensory and causal reasoning abilities necessary for that survival. Humans muddle through complex situations using heuristics and constant readjustments.[16]

A practical approach to the problem of regulating robot behavior is to focus on guidelines and policies for robot manufacturers, distributors, owners, and users, placing accountability and liability with these stakeholders, and not with the robots.

Legal Remedies

Laws, regulations, and policies are based on theories of justice and ethics, social convention and expectations, and the ever-evolving moral (and sometimes immoral) responses of juries, judges, business executives, journalists, and other social influencers to novel circumstances. In common law countries such as the United Kingdom and the United States, the universe of case law expands through analogy—new legal circumstances are typically interpreted through the lens of legal precedent. Legal decisions that involve new technologies, in particular, are often guided by analogies to older, better understood technologies. In the case of robotics, laws, regulations, and policies will evolve according to the ways in which different types of robots are viewed as tools or as autonomous

[16]Hollnagel, E. (1992, March). Coping, coupling and control: the modelling of muddling through. In *Proceedings of 2nd interdisciplinary workshop on mental models* (pp. 61-73).

agents and, according to the unique differences between robots, older forms of technology (such as animal domestication),[17] and other emerging technologies (such as cloning and genetic modification). Just as the Internet challenged privacy, property, and commerce laws that had been based on jurisdiction and geography, robots will also challenge our current notions of liability, foreseeability, and intentionality.[18]

Over the past decade, there has been much legislative creativity regarding the domestic use of drones and autonomous vehicles.[19] In the United States, each state has discretion over transportation within its boundaries, over and above the rules set by the US Department of Transportation (US DOT). In 2011, Nevada became the first state to permit

[17]There is an interesting debate on the extent to which laws concerning autonomous robots should be modeled on laws regarding the legal status of animals and the consequences of their actions. Much of the discussion concerns the autonomy of the animal and whether its actions were foreseeable by the owner. Animals with wild unpredictable histories confer a different level of obligation on owners, but specifics vary from nation to nation, as noted by, Kelley, R., Schaerer, E., Gomez, M., & Nicolescu, M. (2010). Liability in robotics: an international perspective on robots as animals. *Advanced Robotics*, 24(13), 1861-1871. In contrast, it can be argued that the analogy to animals provides little benefit. Although key issues such as foreseeability, experience, training, and control are important for both robots and animals, the differences in how these are achieved limit the value of the analogy; see Johnson, D. G., & Verdicchio, M. (2018). Why robots should not be treated like animals. *Ethics and Information Technology*, 20(4), 291-301.

[18]Calo, R. (2015). Robotics and the Lessons of Cyberlaw. *California Law Review*, 513-563.

[19]Demiridi, E., Kopelias, P., Nathanail, E., & Skabardonis, A. (2018, May). Connected and Autonomous Vehicles–Legal Issues in Greece, Europe and USA. In *The 4th Conference on Sustainable Urban Mobility* (pp. 756-763). Springer, Cham.

autonomous vehicles, and in 2012 several other states passed their own laws. As an example, California's 2012 law[20] set forth conditions for:

- *Testing*: There must be proof of insurance, and an agent of company must sit in the driver's seat, monitor its activities, and be able to take full control of the vehicle in an emergency.
- *General operation*: The vehicle must be insured, meet all safety standards, successfully complete all tests on public roads, and must comply with all state standards.

Of interest to our discussion is that the law defines *operator* as the person in the driver's seat or as the person who initiates the operation of the autonomous technology, and defines the *manufacturer* as the entity or person that equips the vehicle with autonomous technology. The manufacturer must ensure that

- The operator can visually confirm that the autonomous technology is engaged
- The operator can easily disengage the autonomous technology
- If a failure is detected, the system requires the operator to take control or, if that's not possible, to come to a safe, complete stop
- The vehicle contains a "black box" device that records at least 30 seconds of sensory data prior to a collision

[20]Codified into state transportation law as Vehicle Code, DIVISION 16.6—Autonomous Vehicles [38750], https://law.justia.com/codes/california/2012/veh/division-16.6/section-38750/ [accessed on March 5, 2020]. For information on The Federal Automated Vehicles Policy which was initiated in 2017 as a set of recommendations to the states, see www.ncsl.org/research/transportation/regulating-autonomous-vehicles.aspx

We expect this pattern to continue as robotic technology evolves—a clear chain of responsibilities from manufacturers to owners to human operators. The humans that work with and especially those that operate robots will need to understand their obligations for monitoring and at times controlling robotic behavior. Most importantly, under current law, humans or organizations (of humans) are solely responsible. Robots are considered as property with no independent responsibility (or rights), regardless of its inherent intelligence or autonomy from human beings.[21]

For decades, software has played an essential role in our economy, entertainment, legal systems, and daily life. However, in most cases, decision-making that has real-world, irreversible consequences has been the domain of humans. Unlike pure software applications (such as search algorithms or recommendation systems), robots can manipulate things in the world—they can construct buildings, move heavy objects, and cause harm. As robots become integrated into our physical world, there will be a corresponding increase in accidents and injuries involving robots,[22] from malicious intent by humans to use robots to injure other people to intended and unintended harms created by robots.

Are Robots Exceptional?

Robots vary in their degree of autonomy and in how they interact with the world. Robots built for warfare are very different from industrial robots that are meant to operate in isolation, and these in turn are different from cobots that work alongside humans. Moreover, the context in which humans and robots encounter each other vary widely: a person who has

---

[21]Pepito, J. A., Vasquez, B. A., & Locsin, R. C. (2019). Artificial Intelligence and Autonomous Machines: Influences, Consequences, and Dilemmas in Human Care. *Health*, *11*(07), 932.

[22]Kelley, R., Schaerer, E., Gomez, M., & Nicolescu, M. (2010). Liability in robotics: an international perspective on robots as animals. *Advanced Robotics*, *24*(13), 1861-1871.

consented to, and is aware of the risks of, robotic surgery is in very different context from a pedestrian who encounters an autonomous, driverless vehicle.

How does the law consider negligence, obligations, and liability in all these different situations? Can a robot be liable, or should responsibility always be attributed to a human, such as the manufacturer, distributor, owner, or operator? Should we create a new type of law to handle these questions? And, what can businesses do to manage the risk of employing robots?

Legal frameworks must balance competing concerns and rights of individuals and society.[23] On the one hand, they must protect the rights of human consumers and employees, and on the other hand, they must provide incentives and flexibility so that businesses can use and extend robot technology to produce innovations that are desired by consumers.

Legal codes throughout the world generally discriminate between different areas of application, such as criminal, maritime, military, and civil laws. In response to the unique social challenges created by the Internet and autonomous vehicles, some legal scholars have proposed a *strict exceptionalism* in which cyberlaw (i.e., laws regulating the Internet) and robotics are considered new legal domains, each with distinct laws.[24] The argument for strong exceptionalism rests on the belief that the Internet and robotics are transformative technologies that will create (or have created) novel situations that are not handled by current statutory or common law. The Internet, for example, challenges accepted notions of jurisdiction: if a California company's website is accessed in Pennsylvania, can it be sued in Pennsylvania for violating that state's trademark or commerce laws?

---

[23] *Ibid*.

[24] Although the Internet and robotics share many qualities, they present distinct legal challenges. See Calo, R. (2015). Robotics and the Lessons of Cyberlaw. *California Law Review*, 513-563.

In contrast to the Internet, robots have a physical location (although they might access and receive instructions through the Internet). Because of their physical embodiment, robots combine the unconstrained and viral nature of information with the capacity to physically harm people and property. "… Robots, more so than any technology in history, feel to us like social actors—a tendency so strong that soldiers sometimes jeopardize themselves to preserve the 'lives' of military robots in the field."[25]

Strong exceptionalism is a legal perspective that argues that certain technologies or situations create legal conflicts that are so strikingly different from legal precedent that a new legal framework is needed. Some legal scholars, for example, have suggested that just as maritime law has distinct rules and institutions, the Internet and virtual worlds are independent legal entities, "separated from doctrine tied to territorial jurisdictions."[26] However, according to Greg Lastowka and Dan Hunter:

> *The Internet, despite early predictions, never became an independent community. Websites and other prior technologies of cyberspace served as remarkable tools for communication, but they did not build truly independent and self-governing communities. By contrast, avatar existence and avatar community* only *occurs [sic] within virtual worlds, making the emergence of virtual law within those worlds much more likely.*[27]

In contrast to the legal arguments for exceptionalism, many legal scholars argue that exceptionalism is typically the wrong way to teach and evolve law. Just as you would not have a special code of law dedicated

[25]Calo, R. (2015). Robotics and the Lessons of Cyberlaw. *California Law Review*, 513-563.

[26]Johnson, D. R., & Post, D. (1996). Law and borders: The rise of law in cyberspace. *Stanford Law Review*, 1367-1402.

[27]Lastowka, F. G., & Hunter, D. (2004). The laws of the virtual worlds. Calif. L. Rev., 92, 1.

to horses, that is, "The Law of the Horse,"[28] we should not have a Law of the Drone, or a Law of the Internet, or a Law of the Robot. According to this view, society should develop through statutes and case law, sound principles of privacy, property, and liability and then apply it to Internet transactions, robotics, and other new technologies. The result will be deeper, more general legal concepts that are refined through legal challenges, rather than a collection of potentially inconsistent legal rulings. Or, to put it in terms of horses, "Only by putting the law of the horse in the context of broader rules about commercial endeavors could one really understand the *law* about horses."[29]

Some legal scholars, such as Ryan Calo,[30] recommend a moderate version of exceptionalism that is halfway between *strong exceptionalism* and *no exceptionalism*. As he points out, there are already specific statutes dealing explicitly with drones and autonomous vehicles. He argues, however, that new technologies sometimes create imbalances or conflicts that can only be systematically resolved through fundamental changes to law or the introduction of new regulatory institutions. We share this perspective: *Robotics does not present a complete discontinuity with the past—fundamental notions of liability and privacy still apply, for instance. Rather we can make sense of, and adapt, our current legal framework to the evolving robotics industry*. For example, just as the introduction and massive adoption of radio led to a new regulatory institution (what evolved into the US Federal Communications Commission), legal conflicts regarding autonomous robots might motivate the formation of a Federal Robotics Commission.

---

[28] Calo, R. (2015). Robotics and the Lessons of Cyberlaw. *California Law Review*, 513-563.

[29] Easterbrook, F. H. (1996). Cyberspace and the Law of the Horse. *U. Chi. Legal F.*, 207. The quotation maintains the italicized word, "law."

[30] Calo, R. (2015). Robotics and the Lessons of Cyberlaw. *California Law Review*, 513-563.

Can a Robot Be Biased?

There are many different types of robots. At one extreme are robots whose behavior is deterministic—its errors are due to design, manufacturing,[31] or maintenance flaws, or because the operator/owner did not read the warning labels and used the robot incorrectly. For example, if the product warnings indicated, "dry use only," and the owner placed the vacuum in an outdoor pool, the ensuing damage would likely be the owner's responsibility.

At the other extreme are robots whose behaviors are based on in vivo deep learning or other stochastic processes that are not deterministic—each response to a given situation might not be foreseeable by its designer, manufacturer, software programmer, trainer, owner, or operator. An inappropriate response could be based on unforeseeable experiences that occurred after product delivery and training, and therefore not under the control of the agents typically involved in product production and consumption. For example, let's assume that a robot is trained as a greeter in a bank. During its training phase, the robot asked each bank visitor their name and how they were feeling. During the week of training, most of the bank visitors just happened to be white men from North America. As a result, the robot's face and affect recognition abilities were biased and

[31]Robot design and manufacturing flaws can be introduced by hardware, firmware, or software providers. Design defects affect all the instances of a product—the product is manufactured as intended. Design defects reflect conscious choices by the manufacturing, although with unintended consequences. In contrast, manufacturing flaws occur randomly and are not detected during standard and reasonable quality inspections. They may be introduced though flawed materials or during production and may be limited to a single instance or to only a those produced during a production run. See Tietz, G. F. (1993). Strict products liability, design defects and corporate decision-making: greater deterrence through stricter process. *Vill. L. Rev.*, *38*, 1361.

some visitors who do not meet that description were ignored or treated incorrectly.[32]

Who is responsible for this bias and the resulting indignities? Possible answers might be the software manufacturer, the company that oversaw training, the company that oversees maintenance and quality assurance, or the bank and its employees who are trained on proper care and monitoring of the robot.

Who Is Responsible?

Under current law, robots and the software/firmware that direct their behavior are commercial products and considered property—they have human owners. They are introduced and traded in the marketplace by humans. Humans bear responsibility and liability for these commercial transactions.[33] In most US jurisdictions, manufacturers or sellers of a product can be strictly liable for harm caused by design defects, manufacturing flaws, and inadequate information about proper use (e.g., warning labels).[34]

[32] This example is not a fanciful construction. There are many real-world examples, of bias in machine-learning algorithms and training data. In some case, the bias is intentionally created through product use by hate groups. The book, *Algorithms of Oppression* (Noble, S. U., 2018) recounts, for example, how the actions of white supremacist groups influence Google's search results for the search term, "Jew," and how unmoderated racist discussions bias the results for "black girls" vs. "white girls." Microsoft's chatbot, *Tay*, was intentionally corrupted by hate groups creating racist dialogue during *its* in vivo training. But bias can also be created unintentionally through sampling errors, as in the example of the bank's robot.

[33] Johnson, D. G., & Verdicchio, M. (2018). Why robots should not be treated like animals. *Ethics and Information Technology*, *20*(4), 291-301.

[34] The application of strict liability for design defects and warnings may vary from state to state. Some states apply a consumer expectations test in which allows plaintiffs to argue that the product was unsafe for reasonably foreseeable uses and abuses (in addition to its intended use). Alternatively, a risk-utility test allows defendant to argue that no alternative design could have reduced the foreseeable harm and maintained the product benefits; see Abeyratne, R. (2017). Artificial Intelligence and Air Transport. In *Megatrends and Air Transport* (pp. 173-200). Springer, Cham.

However, assuming no product defect or that the owners assume the risk of defect, then the owners might be liable if they are negligent in their duty of care to those who might be affected by the robot's behavior. This might be caused by neglecting proper maintenance, training, or usage.

An injured person might bear proportional responsibility if they interacted with the robot in a dangerous or intentionally abusive manner, for example, if someone knowingly made themselves difficult to detect and entered the predictable path of a robot, or (as mentioned earlier) if they jumped over a security fence, they should bear some responsibility for the consequences.

Under US tort law, a malfeasant could be liable for any of the harms to other humans and property if their actions generated significant foreseeable danger[35] as could happen during the beta test of a collaborative robot.

According to the European Union directive on product liability, manufacturers must assume responsibility for assuring "that its products are suitable for their intended use when they are placed on the market."[36] However, robotics technology will likely complicate the definition of manufacturer. Manufacturers are the person or business that places their name or trademark on the product, or the importer of the product, or any person supplying the product in a transaction.[37] However, in the case of robots, we should differentiate between:

1. *The hardware manufacturer* who produces or integrates components into an independent, movable, unified object that can interact with its environment through sensors and actuators.

---

[35]Shavell, S. (2018). The Mistaken Restriction of Strict Liability to Uncommon Activities. *Journal of Legal Analysis, 10*.

[36]Pepito, J. A., Vasquez, B. A., & Locsin, R. C. (2019). Artificial Intelligence and Autonomous Machines: Influences, Consequences, and Dilemmas in Human Care. *Health, 11*(07), 932.

[37]Zornoza, A., Moreno, J. C., Guzmán, J. L., Rodríguez, F., & Sánchez-Hermosilla, J. (2017). Robots Liability: A Use Case and a Potential Solution. *Robotics: Legal, Ethical and Socioeconomic Impacts*, 57.

2. *The software manufacturer or programmer* who provides the logical apparatus for storing and discarding information, image processing, decision-making, and so on.

3. *The data provider/trainer* who provides any data and training required prior to the original sale of the robot.

4. *The seller of the robot*, usually one or more of the preceding three categories of manufacturers, who assumes strict liability for design, data, production, and marketing defects in the original product transaction.

5. *The owner* of the robot (typically a business entity).

6. *The user or operator* of the robot (typically an agent of the owner of commercial robots). Notably, the operator will likely be trained in how to operate and shut down the robot, as well as in a specialty, such nursing, search-and-rescue, warehouse operations, logistics, trucking, and so on.

7. *Bystanders* who may or may not be expected to behave in certain ways and have the necessary knowledge or experience.

Lastly, as artificial intelligence technologies allow more autonomy, perhaps we should also consider the culpability of the robots, themselves and any collective machine learning that is aggregated and redistributed through the cloud. When deep-learning systems are used, the results are not always foreseeable. The robot system may learn the wrong things, such as misclassifying a pedestrian carrying a flashlight as a threat, because in the training data most intruders carried a flashlight. Or, a random choice is made in the robot's classification decision and the downstream effect is denying parole to a convict. Is there a point at which we will need to consider autonomous robots as something different than property and hold them liable?

In this chapter we have taken the view that intelligent, autonomous machines should not be considered moral agents, at least for the time being. They are property and their manufacturers, owners, and operators have moral obligations to provide and maintain robots that behave safely in all foreseeable circumstances.

Robots in Corporations, Corporations in Robots

> *Participation of users in the design of robots can also allow multiple perspectives on technology and society to be expressed in the course of deciding on the uses and technological capabilities of robotic artifacts. The explicit and systematic exploration of the feedback between social and technological choices can inspire reflection by robot designers, analysts, and users on the social norms and values robots embody and enable us to mindfully create more socially robust, responsive, and responsible robots.* [38]
>
> —Šabanović, S. (2010)

Modern software engineering is guided by best practices, such as agile development and engineering practices,[39] and each product is defined by

[38]Šabanović, S. (2010). Robots in society, society in robots. *International Journal of Social Robotics*, *2*(4), 439-450.

[39]Abrahamsson, P., Salo, O., Ronkainen, J., & Warsta, J. (2017). Agile software development methods: Review and analysis. *arXiv preprint arXiv:1709.08439*.

its functional and nonfunctional requirements. Functional requirements define the capabilities or features of a product. The nonfunctional requirements define attributes such as scalability, usability, reliability, performance, and privacy. Although they may not be directly reflected in the user stories, features, or capabilities of a product, nonfunctional requirements strongly influence system architecture and act as constraints on how features or capabilities are designed. They may not be achievable in the first prototype, or the first minimally viable version of a product, but professional software and hardware architects are cognizant of these constrains from the beginning of product development.

In this subsection, we argue that ethical constraints, including privacy, should be integrated into software design and development practices. We begin by discussing privacy as a nonfunctional requirement that needs to be considered in the initial phase of product design.

Privacy by Design

Ann Cavoukian, a seminal privacy rights advocate, was Ontario, Canada's Information and Privacy Commissioner, from 1997 to 2014. In 2010, her *privacy by design* concepts achieved international recognition.[40]

Privacy by design is an engineering practice that provides a framework for assuring that data and software applications maintain reasonable levels of privacy. Cavoukian's radical idea was that security and privacy are not a zero-sum game. Rather, the legitimate objectives of both can be accommodated, especially if privacy is incorporated into the technology and architecture and not added on as a post hoc business practice.

[40]Landmark Resolution passed to preserve the Future of Privacy (2010). www.ipc.on.ca/english/Resources/News-Releases/News-Releases-Summary/?id=992 [accessed on February 21, 2020].

Privacy by design stipulates seven basic systems engineering principles:

1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default
3. Privacy Embedded into Design
4. Full Functionality—Positive-Sum, Not Zero-Sum
5. End-to-End Security—Lifecycle Protection
6. Visibility and Transparency
7. Respect for User Privacy

Although these were designed to proactively embed privacy into IT design and business practice, the principles can be abstracted and applied as a framework for ethical governance. These will be reflected in the next section.

Privacy by design is incorporated into the EU General Data Protection Regulation (GDPR) that was adopted by the European Parliament and the Council of the European Union in April 2016 and enforced as law in May 2018. In contrast, as a leading indicator of what the United States might do, the California Consumer Privacy Act, which became effective on January 1, 2020, does not mandate privacy by design principles.[41]

Ethics by Design

Nonfunctional requirements such as privacy, ease of use for all users (including those with disabilities), fairness and lack of bias, and the welfare of the affected communities are ethical requirements. Security, reliability and performance requirements protect the integrity of a software system;

[41]OneTrust (December 19, 2019). The CCPA vs. the GDPR. www.onetrust.com/the-ccpa-vs-the-gdpr/ [accessed on February 21, 2020].

ethical requirements protect the welfare of the people and the community that contains and interacts with the system.

AI ethics is a matter of significant concern. The public, AI and legal professionals, and businesses are all rightfully concerned about ethical errors in judgment, bias, and liability that might occur with the application of AI-enabled robots and automation. An insightful paper by Hagendorff, concludes that although much effort has been invested in developing guidelines, "… the practice of development, implementation, and use of AI applications has very often little to do with the values and principles postulated by ethics."[42]

There are many well-documented failures, but there is no clear evidence that ethical guidelines are intentionally incorporated into decision-making practice by developers, software manufacturers and retailers, and the software itself. In short, ethics by design is at the same stage that privacy by design was at in the mid-1990s: concepts embraced by a well-meaning professional community but unenforced by technology and with uncertain legal consequences (unless there is a specific violation of existing law).

Nonetheless, it is worthwhile to explore how ethics by design might positively impact robot design and how that impact might be realized through best practices and law. The objective is to reduce the risk and to optimize the benefit for the well-being of the community, the enterprises that build and distribute robotic systems, and its workers.

[42]Hagendorff, T. (2019). The ethics of AI ethics—an evaluation of guidelines. *arXiv preprint arXiv:1903.03425*.

Recent discussions on AI ethics focus on[43]

- Embedding ethics into machine-learning and decision-making algorithms, including specific problems in specific domains (such as decisions by autonomous vehicles that will injure its passengers or, instead, injure nearby pedestrians)[44]
- Meta-analytic or expert-based derivations of general principles to guide design, development, and deployment

Hagendorff (2019)[45] systematically reviewed 21 of the most influential ethical guidelines for building AI systems proposed by academics, nonprofits, and industry. This review identified 22 characteristics; many of which were shared across guidelines. For example, accountability, privacy, and fairness each appear in 17 out of 21 guidelines, and transparency/openness, cybersecurity/safety, and common good appear in 15 of the 21 guidelines.

Hagendorff warns that the lists have biases. The most mentioned characteristics are already the focus of industrial and academic research. He also notes that most of the guidelines are authored primarily by men (roughly two-thirds of the collective authorship) and that most of these analyses tend to focus on characteristics that can be isolated, transactional, and defined as technical problems with technical solutions. In contrast,

[43]*Ibid*.

[44]Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self-driving cars: An applied trolley problem? *Ethical theory and moral practice*, 19(5), 1275-1289.

[45]Hagendorff, T. (2019). The ethics of AI ethics—an evaluation of guidelines. *arXiv preprint arXiv:1903.03425*.

the reports of *AI Now*[46] (in which women were the primary co-authors) do not frame "AI applications in isolation, but within a larger network of social and ecological dependencies and relationships,"[47] with an emphasis on the ethics of care and social well-being. This difference in emphasis underscores the importance of gender, ethnic, and cultural diversity in the definition of any ethics guideline.

The 22 characteristics identified by Hagendorff can be clustered into five major categories, as shown in Table 7-1.[48] These five categories generalize the *privacy by design* principles. The categories are not meant to be mutually exclusive perspectives, rather they inform one another. The number in the brackets identifies how many guidelines mentioned that characteristic.

[46]Whittaker (2018).

[47]Hagendorff, T. (2019). The ethics of AI ethics—an evaluation of guidelines. *arXiv preprint arXiv:1903.03425*.

[48]This clustering was not provided by Hagendorff; we found it is useful for the present discussions and any flaws are the responsibility of the present authors.

Table 7-1. *Ethics by Design Categories*

Categories	***Characteristics identified in published AI Ethical Guidelines***
1. Design representative of diversity	• Diversity in the field of AI [6] • Cultural differences in the ethically aligned design of AI systems [2]
2. Accountability, explainability, and transparency	• Accountability [17) • Transparency, openness [15] • Human oversight, control, auditing [12] • Explainability, interpretability [10] • Legislative framework, legal status of AI systems [9] • Certification for AI products [4]
3. Governance	• Science-policy link [10] • Responsible/intensified research funding [8] • Public awareness, education about AI and its risks [8] • Field-specific deliberations (health, military, mobility, etc.) [7] • Protection of whistleblowers [2)
4. Safety	• Privacy protection [17] • Safety, cybersecurity [15] • Dual-use problem, military, AI arms race [7]
5. Social Impact—well-being	• Fairness, nondiscrimination, justice [17] • Common good, sustainability, well-being [15] • Solidarity, inclusion, social cohesion [10] • Future of employment [8] • Human autonomy [7] • Hidden costs (e.g., environmental and energy costs) [1]

The number in the brackets identifies how often that characteristic was identified in the AI Ethics Guidelines reviewed by Hagendorff (2019).

The next five subsections discuss the significance of each of the five categories identified in Table 7-1.

Design Representative of Diversity

Who will the robot interact with, directly and/or indirectly, during its design, development, and use? AI systems and robots in particular may affect many different kinds of people varying in gender and sexual orientation, racial, cultural, linguistic and ethnic backgrounds, gender, cognitive and physical abilities, socioeconomic status, and age.

The robotic system must not discriminate against any of the people affected by the system during its design, development, and deployment. This doesn't mean that it treats all users in the same way, but rather that it treats all stakeholders (programmers, trainers, users, and bystanders) appropriately and fairly—it may need to perform its tasks differently for the elderly and young, for the physically challenged, and so on, but must do so in a way that respects their dignity and supports their social wellness. To accomplish this the design must be:

Proactive, preventative, and adaptive: This is important in all AI systems, but it is especially true in robotics where nonreversible actions on the physical world can occur. The systems must be designed to interact with and anticipate the needs and challenges of all user classes. If a face recognition system is used, it must perform well, for example, across all racial and age groups.

Designing for all classes of user must be mindful of intersectionality.[49] Designing a system that is fair and appropriate for all genders and for all ages must consider the intersection (or interaction) between these—designing a robotic system for use by a young boy is different than simply considering the individual needs of children and the needs of males; the way in which women, minorities, and the disabled are treated within the

[49]Crenshaw, K. (1990). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stan. L. Rev.*, *43*, 1241.

US culture is different and not entirely predictive of how minority women with disabilities are treated. Designers must carefully consider the different groups that are affected by their designs.

Ethics embedded into design, development, and deployment: One of the key approaches to reducing harm to a demographic group is having representatives of that group directly involved in the design, development, testing, and deployment of the product. Many AI systems and robots are developed without careful ethnographic studies of the current practices that they are seeking to replace or transform. Unfortunately, many of the systems are designed, developed, and tested by a homogenous group of 20-to 30-year-old white males. For *Ethics by Design*, there should be more woman and minority representation, and age diversity, in the AI and robotics workforce, at all levels of management and decision-making.[50]

In addition, care should be taken to examine the social impact of these robots. Will they affect jobs across all industries, or will they create more benefit or harm to certain communities? Representatives from various affected communities should be consulted throughout the product life cycle. This will improve the benefits, decrease the harms, and increase community acceptance.

Accountability, Explainability, and Transparency

In 2016, ProPublica provided strong statistical evidence that Northpointe's software was biased in their recommendations to sentencing judges about

[50]The negative impact of software designed by an ethnically homogenous group is well documented in Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press; and Chang, E. (2019). *Brotopia: Breaking Up the Boys' Club of Silicon Valley*. Portfolio.

the likelihood of recidivism.[51] The manufacturer claimed that the statistical evidence was misinterpreted, but they refused to disclose details about the algorithm or data handling processes, arguing that the algorithm and data were competitive advantage and proprietary.

Under what conditions should manufacturers or trainers (of machine-learning systems) be required to disclose the data and algorithms used in a decision? Should this only happen after demonstrated bias or a tragic outcome? When someone is declined a loan, they should be provided an explanation for the denial, but should they be given details about the algorithm?

We have already argued that a human or organization (e.g., a corporation or government) should always be liable for the actions of their robotic systems; robots should never be considered inscrutable, independent moral agents. When harm occurs or an undesirable decision is made, those who are negatively affected should be able to know the causes.

One approach to understanding the cause of a decision or action is to construct algorithms that are *self-explaining*—they provide explanations for their behavior either in real-time or upon request after the action has occurred. Another approach is to require *transparency*—manufacturers would disclose their algorithm and training data prior to deployment in order to receive certification that ethical standards have been met, or they would be required to allow forensic inspection of their software and data, by court order.

[51] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. *And it's biased against blacks. ProPublica, 23.* Northpointe and others have disputed ProPublica's findings; see, for example, Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). False Positives, False Negatives, and False Analyses: A Rejoinder to Machine Bias: There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks. *Fed. Probation*, *80*, 38.

Explainable AI can be defined as the methods and policies that allow the results of an algorithm to be explained by experts or even by another AI program. The explanation may be the output of a forensic investigation to examine a system failure or bias, or it can be delivered as part of the initial decision, as in "Loan approval is recommended for the following reasons …." One common method for achieving explainability is to use rule-based algorithms which can report the weight or risk value associated with each rule used in the decision. If self-reporting is not possible, data analysts can attempt to examine the data and deep-learning system (layer by layer, region by region) to discover what data were used and what impact that had on the results.

However, whether the algorithm is explainable automatically or through rigorous inspection by the manufacturer, the explanation must be verifiable by independent agencies. Some form of transparency is required. The challenge is that the data and algorithm are intellectual property of the manufacturer. If manufacturers can be compelled to expose data and code to public scrutiny, they might lose the competitive advantages gained by years of investment and research.

One proposal is to form public governance boards that can certify or inspect confidential algorithms and databases.[52] These boards would operate independently of the manufacturer that created the algorithm but would uphold confidentiality. In order to conduct a reasonable audit, the board could review the source code and the data provided during training and during its deployment. However, this is not only tedious, but static methods provide little insight into how machine-learning algorithms interact and adapt to data or to its real-time environment. Conducting dynamic analyses on the machine-learning algorithm after it has been

[52]The AI Now report (2018) recommends, "Governments need to regulate AI by expanding the powers of sector-specific agencies to oversee, audit, and monitor these technologies by domain." Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., … & Schwartz, O. (2018). *AI now report 2018*. AI Now Institute at New York University.

trained but before it's deployed also has limits. The test data will be a subset of the potential interactions with the environment. If the machine-learning process utilizes randomization in its training or application, repeatability is not possible—the same data inputs will lead to different outcomes because of random differences in the outcome of intermediate algorithmic decisions. However, for forensic purposes, software systems can store the randomization that were used in vivo, and these can be reused when the algorithm is analyzed. Consistent with this approach, some researchers have suggested a *black box* requirement in which robots and autonomous systems are required to store all the relevant data (sensory data, interim decisions, etc.) in a write-only, secure repository.[53]

Manufacturers are likely to resist requests to share their code with government or third-party reviewers. One proposal is to build accountability into the software architecture, so that indicators of bias, risks, and other influences can be analyzed forensically.[54] This would enable review boards to operate without having manufacturers fully expose confidential data and algorithms.

In response to corporate reluctance to partly or fully expose their algorithms and data, legal scholars have argued that existing legal doctrine should be used to make vendors of AI and robotic systems more accountable, especially in systems used by the government. As Crawford and Schultz (2018) note, "…as AI systems rely more on deep learning, potentially becoming more autonomous and inscrutable, the accountability gap for constitutional violations threatens to become broader and deeper."[55] The blame for violations of policies or law must

---

[53]Winfield, A. F., & Jirotka, M. (2017, July). The case for an ethical black box. In *Annual Conference Towards Autonomous Robotic Systems* (pp. 262-273). Springer, Cham.

[54]Kroll, J. A., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2016). Accountable algorithms. *U. Pa. L. Rev.*, *165*, 633.

[55]Crawford, K., & Schultz, J. (2019). AI SYSTEMS AS STATE ACTORS. *Columbia Law Review*, *119*(7), 1941-1972.

target manufacturers and data providers. *Blame should never apply to some inscrutable algorithm that no one fully understands, and that plaintiffs cannot inspect or hold financially accountable.*

Accountability laws and policies have implications for the workforce. As noted by Wilson et al.,[56] machine learning will create many new job categories including ethics compliance managers that focus on automated decision-making. These managers will need advanced certifications and multiple skill sets. They will be tasked with evaluating and monitoring robotics systems during design, development, and after deployment. They will have to be able to explain to plaintiffs, lawyers, and juries why certain behaviors were exhibited and whether certain biases or errors were foreseeable. They will also need to be skilled at generating edge cases for testing various scenarios.

Governance

Industry, health services, and governments (including the military) are rapidly expanding the use of robots and automation in their decision-making. Robots can be, or will soon be, involved in manufacturing, transportation, surgery and diagnostics, surveillance, search-and-rescue, and warfare. Some businesses are installing procedures and oversight to assure transparency and auditability, and the input of stakeholders, but many are not. The robotics and AI industry needs government and nonprofits that can develop, certify, and enforce standards compliance and well-defined policies.[57] This has worked reasonably well in other industries, such as food safety which is monitored through government

[56]Wilson, H. J., Daugherty, P., & Bianzino, N. (2017). The jobs that artificial intelligence will create. *MIT Sloan Management Review*, *58*(4), 14.

[57]Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., … & Schwartz, O. (2018). *AI now report 2018*. AI Now Institute at New York University.

agencies (e.g., the US Food and Drug Administration) and through nonprofits (e.g., Green Seal, a nonprofit that certifies products and services that meet health and environmental impact standards).

The 2018 *AI Now Report* recommends that governments should expand sector-specific nonmilitary agencies to oversee, audit, and monitor AI technologies.[58] We agree with this approach and recommend that it be expanded to include robotics, including military robots. These efforts must balance the need for innovation with the need for worker and public safety.

Governments should also apply pressure through existing legal doctrine, to encourage companies to disclose when AI is used in automated processes, such as loan decisions. Consumers and employees respond well when companies are transparent about how data is used and what ethical guidelines they follow.[59]

This need for governance though public and nonprofit organization suggests that corporate ethics committees should be created at the C-suite level, and not within the IT department. To be effective they will need to work with suppliers, unions, and employees and with communities and consumers. We expect that many corporations will soon designate Chief Ethics Officers who will oversee the incorporation of ethical principles into workplace robots and automation, in addition to other responsibilities.[60]

---

[58]Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., ... & Schwartz, O. (2018). *AI now report 2018*. AI Now Institute at New York University;

[59]Clement-Jones, T.L., joined by Thomas E. (September 3, 2019). Future regulation for artificial intelligence. *DLA Piper TechLaw Podcast Series* [audio accessed on February 6, 2020].

[60]Insights Team (March 27, 2019). Rise of the Chief Ethics Officer. *Forbes Insights*. www.forbes.com/sites/insights-intelai/2019/03/27/rise-of-the-chief-ethics-officer/#5b1f5c105aba [accessed on April 3, 2020].

Safety

> *Silicon Valley is known for its "move fast and break things" mentality, whereby companies are pushed to experiment with new technologies quickly and without much regard for the impact of failures, including who bears the risk.*[61]
>
> —Whittaker, M., *et al*. (2018)

With the exception of autonomous vehicles where state regulations must be followed, AI systems are deployed without consulting with communities and workers that are directly or indirectly impacted. Workers have good insights into the errors and unsafe conditions that might occur within their environment and their participation in design can promote innovations that create safe technologies and improved processes.[62]

Governance regarding safe operations should be automated through proper onboard and external controls. Robotic systems should shut down systems or revert to a safe configuration when a robot behaves out of bounds, for example, entering an area with humans. Likewise, robots and AI programs that attempt to evaluate the affect or health of a human should not perform a nonreversible decision or action without human supervision. The likelihood of false classifications and the inflexibility of the consequences, in today's systems, suggests the need for careful governance and not efficiency. For example, a police drone should not apprehend or question pedestrians if they look suspicious, but rather should maintain privacy standards and log its concerns with nearby police officers.[63]

---

[61] Whittaker, M., et al. (2018). AI now report 2018. AI Now Institute at New York University.

[62] See Biddle, R., Brown, J. M., & Greenspan, S. (2017). From Incident to Insight: Incident Responders and Software Innovation. *IEEE Software*, *36*(1), 56-62, for a careful discussion on how operations staff can help product designers.

[63] McNeal, G. S., Goodwin, W., & Jones, S. (2017). Warrantless Operations of Public Use Drones: Considerations for Government Agencies. *Fordham Urb. LJ*, *44*, 703.

As a society we have moved from pedestrian policeman patrolling a neighborhood to police cars navigating a neighborhood, to webcams surveilling a neighborhood, and now to semi-autonomous drones and other robots traversing and protecting the neighborhood,[64] guided by algorithms that predict where the next crimes are most likely to occur. With each step there are efficiency gains—fewer police can do more, but in the process they may lose core skills,[65] increase the distrust between police and community, and decrease the safety of some communities (in particular, minority or poor). We should learn from this history—industry, government, and local communities should work together to create safe and effective applications of robots.

The concerns for safety are especially noteworthy in military applications. The *Campaign to Stop Killer Robots* is a coalition of NGOs formed in 2012 to persuade the UN to adopt a legally binding ban on fully autonomous weapons. The push for such weapons is based on myths suggesting that autonomous weapons will not tire or act irrationally, they won't rape or seek revenge, they will obey the conventions of war, and they will save soldiers' lives and kill fewer civilians.[66] However, as we have argued earlier, autonomous robots make mistakes, just a humans do, but they do it faster, with greater precision and at scale. Their errors can be more difficult to stop. What if the wireless signal intended to stop the autonomous weapon is jammed and not received? They may not be able to discriminate between soldiers and civilians; they may be used in ways

[64]Stephen Rice (October 7, 2019). 10 Ways That Police Use Drones To Protect And Serve. *Forbes*. forbes.com/sites/stephenrice1/2019/10/07/10-ways-that-police-use-drones-to-protect-and-serve/#4bc4688e6580 [accessed on April 3, 2019].

[65]Joh, E. E. (2019). The Consequences of Automating and Deskilling the Police. *UCLA L. Rev. Discourse (2019 Forthcoming)*.

[66]N. Sharkey (November, 2017). Killer Robots: The Race for Autonomous Weapons. New Internationalist. https://newint.org/sections/agenda/2013/10/01/killer-robots [accessed on March 10, 2019].

that circumvent conventions and ethics—robots are unlikely to refuse an order or inform on a nation that engages in illegal or unethical behavior. As in all robot applications, but with far greater urgency, humans must have complete situational awareness of how and why military robots are used and must have full governance and control over their real-time activities.[67]

Social Impact and Well-Being

Installing new automation or robotic processes can affect workplace morale and well-being. As noted in Chapter 1, "Will Robots Replace You?", the key issue for the original Luddite movement was not the new technology, per se, but rather the manner in which the technology was used and incorporated into the workplace.

The Luddites were skilled, highly paid workers who were proud of their products and who were angered that the new technologies (1) lowered the quality of the textiles, (2) replaced skilled with less-skilled labor, (3) applied without consulting the skilled labor, and (4) increased the power asymmetry between management and labor. The same issues concern us today. Robotics scientists, engineers, and experience designers should understand and report the potential risks, harms, and benefits of installing a robotic system into the workplace or consumer market. These reports should address questions such as: How does it affect the current workers? Are the current workers expected to train their replacements? What happens when robots and humans interact—are there physical dangers that must be avoided? What happens if the robot is hacked into, or misbehaves? Is there a potential bias in the robot's decision-making or behavior that could introduce power and benefit asymmetries into the workplace?

[67] Sharkey, A. (2019). Autonomous weapons systems, killer robots and human dignity. *Ethics and Information Technology*, *21*(2), 75-87.

Within this context, we need guidelines and protections for whistleblowers, employee unions, or other forms of organizing, and grassroots community coordination with employees. *The intent is not to halt or delay innovation, but rather to guide innovation choices so that industry, workers, and the community are all protected and strengthened—a win-win scenario.*

Summary and Conclusion

In this chapter we considered the complexity of robotic interactions in an unpredictable social environment and some of the ways in which humans could be harmed, physically, emotionally, and legally (such as privacy violations and bias). The harm might be due to design, development, training, or communication flaws or because of erroneous or imprecise data. Humans make mistakes and so will robots.

When human manufacturers err or when service providers cause damage, there are well-defined legal remedies that can be adjudicated through the courts. When a robot makes an error, who is to blame? In this chapter, we have taken the position shared by many others in the legal and AI community that robots are products and property. As with other categories of products, there is a standard of responsibility and liability, from manufacturers and owners to trainers and operators.

However, legal statutes, regulations, and standards are needed to promote ethics in the design, development, and deployment of robotic systems. It is highly unlikely that robots can be embedded with an algorithmic, general model of ethics. Instead, the practices adopted by industry and enforced by law must promote ethics. Today, AI and robot designer and developers are largely from a single demographic group. To create ethical robotic systems, tomorrow's workforce needs to be far more diverse.

In order to fairly resolve legal disputes involving the actions of robots and other automated processes, the underlying algorithms need to be more transparent or designed to be more auditable. Legislative frameworks must balance the desire for innovation with the need for safety and ethical decision-making, but these need not be viewed as a zero-sum struggle. Safety and ethical decision-making principles must be designed into the algorithms and data pipelines from the start, and not tacked on at the end of the manufacturing process.

Governance through legislation, certification processes, and internal corporate controls will emerge as the industry matures. In preparation and to facilitate the processes, governments should promote public awareness of AI and robotics. Lastly, the social impact of robotic applications needs to be studied more extensively. As society evolves to more sophisticated and powerful machine-learning algorithms and robots, industry, government, academia, and communities need to work together to promote human safety and autonomy, and fairness in the decisions that affect them. Software engineering practices and legal frameworks must evolve to support these ethical goals.

CHAPTER 8

Work in the Future

A Summary and Conclusion

> *There has never been, nor will there ever be, a technological innovation that moves us away from the essential problems of human nature ... When we rely exclusively on computation for answers to complex social issues, we are relying on artificial unintelligence.*
>
> —Brousard, M. (2018). Artificial Unintelligence: How Computers Misunderstand the World. MIT Press.

As we conclude this book, the world is beset by a crisis of historic proportions. COVID-19, the disease caused by the SARS-CoV-2 virus, has rapidly spread throughout the globe, leaving death, sorrow, and economic hardship in its wake. Medical professionals and scientists the world over are trying to use their expertise to help monitor, make sense of, diagnose, prevent, and treat this infection.

The first indications of the epidemic were raised by AI software. Early in the morning on December 31, 2019, BlueDot's outbreak risk software alerted its customers that a cluster of pneumonia cases had been reported in Wuhan, China.[1] Other AI services that quantify the risk of infectious diseases noticed this anomaly as well, for example,

[1] www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/.

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1_8

HealthMap and Metabiota.[2] These algorithms use natural language processing algorithms to monitor news and government reports and utilize "global air travel patterns around transit hubs, livestock health reports, among other sources to estimate risk."[3]

On January 9, 2020, the World Health Organization issued its first notice, "Chinese authorities have made a preliminary determination of a novel (or new) coronavirus, identified in a hospitalized person with pneumonia in Wuhan."[4] The technology exists to identify clusters of infections, to validate these identifications through additional monitoring and blood tests, to quickly ascertain travel vectors of all of the inhabitants flying from nearby airports, and to monitor the destinations for similar symptoms. The problem is not technology. The issues are public policy, privacy, and cooperation among national and international jurisdictions. The technical issues having to do with data quantity and quality can be reasonably solved; the social issues require international cooperation, political will, trust, and money.

Despite the obstacles, the medical and scientific community have shown great courage and dedication in tracking the epidemic and in providing assistance and advice throughout the world. In addition to detecting and predicting the spread of infection, AI and robotics provide a spectrum of potential applications that could be used to predict, diagnose, and mitigate the impact of infectious diseases and other massive social

[2] www.technologyreview.com/2020/03/12/905352/ai-could-help-with-the-next-pandemicbut-not-with-this-one/

[3] Inn, T. L. (2020). Smart City Technologies Take on COVID-19. *World Health.*

[4] www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china

disruptions (such as famine).[5] These applications can be roughly classified into five main categories as shown with examples in Table 8-1.[6]

Table 8-1. *Application Categories*

Application Categories	Examples

1. Monitoring, detection, and analytics	• Monitor communication and information flows
	  • Identify and validate useful information and curb the spread of misinformation
	• Monitor news sources and the flow of people and animals
	  • Monitor and predict disease transmission vectors

2. Clinical care	• Diagnosis and screening
	  • Automate the processing and distribution of patient data
	  • Automate blood tests
	• Disease prevention
	  • Decontaminate and clean infected surfaces, clothing and bedding
	• Patient care and disease management
	  • Provide bedside care to hospital and remote patients

(*continued*)

[5] Bullock, J., Pham, K. H., Lam, C. S. N., & Luengo-Oroz, M. (2020). Mapping the Landscape of Artificial Intelligence Applications against COVID-19. *arXiv preprint arXiv:2003.11336*.

[6] Yang, G. Z., Nelson, B. J., Murphy, R. R., Choset, H., Christensen, H., Collins, S. H., … & Kragic, D. (2020). Combating COVID-19—The role of robotics in managing public health and infectious diseases.

Table 8-1. (*continued*)

Application Categories	Examples

3. Logistics and communication	• Optimize communication flows
	  • Use chatbots and RPA to provide public access to health services and to automate fulfillment of those services
	• Autonomous transport services
	  • Transport infected or possibly infected individuals to care facility
	  • Transport contaminated specimens and wastes

4. Continuity of work and maintenance of socioeconomic functions	• Teleoperation and Automation
	  • Continue manufacturing and utility operations through robots and remote control
	  • Provide automation to order supplies, and robots to delivery and restock them in local stores

Some of the applications such as using social robots to ease social isolation and to automate the processing of new patient data are in use today.[7] The pandemic has fast-tracked the introduction of these technologies. For example, the Connecticut-based Maplewood Senior Living facility has introduced robots to help residents maintain social distancing and isolation, and at the Mater Misericordiae University Hospital in Dublin, a pilot RPA project speeds COVID-19 test results, "enabling staff to quickly put infection prevention and control measures in place where necessary."[8]

[7]Developed by the Dublin unit of UiPath Inc., the robotic software application distribute test results from the on-site lab in minutes, "enabling staff to quickly put infection prevention and control measures in place where necessary." Loten, Angus (April 6 2020). *Wall Street Journal* (Online) [New York, N.Y].

[8]*Ibid.*

However, automating venipuncture and subsequent blood analysis is a research challenge, but a solution is currently under assessment for use with humans.[9] If successful, automated or robotic methods for drawing and then immediately test blood samples would both protect health works and greatly facilitate screening. Applications such as this one may someday transform healthcare and disease control.

The Transformation of Work

With increasing regularity AI, automation and robotic systems are transforming the way we work and play. They are altering our expectations about what humans and machines can achieve. These technologies enable us to discover correlations in data and thereby discover new pharmaceuticals or new uses for existing pharmaceuticals, to conduct thousands of experiments in parallel, to have greater success in search-and-rescue missions, and to explore planets through semi-autonomous rovers and satellites. Its impact on surgery, job screening, and customer care is more complex with both positive and negative outcomes. And, the potential of military robots is frightening.

AI and robotics are the result of deep yearnings within society and humanity for help—someone or something that can provide wise guidance or accomplish tasks that are too difficult, dangerous, or undesirable. We have also long known about the dark side of these yearnings: the danger of a malevolent superintelligence, the damage caused by an out-of-control wish to a Genie, or the slow self-destruction created by too much idleness (because others are doing the work and making the decisions).

[9]See, for example, Leipheimer, J. M., Balter, M. L., Chen, A. I., Pantin, E. J., Davidovich, A. E., Labazzo, K. S., & Yarmush, M. L. (2019). First-in-human evaluation of a hand-held automated venipuncture device for rapid venous blood draws. *Technology*, *7*(03n04), 98-107.

In this book, we have taken the middle path, examining the benefits, disruptions, and misfortunes, but with the conviction that proper diligence and human governance can create a better society in which tasks that are too dangerous, difficult, dull, or dirty are done by robots with human guidance; that by applying ethics-by-design principles, manufacturers can design and develop collaborative robots that operate alongside, symbiotically, with humans.

Artificial Unintelligence

When discussing the potential of robots and automation to transform work and culture, the question of artificial intelligence is often raised: Will machines become as intelligent as we are? Or, more intelligent? Will they take over the world and rule humanity? And how soon?

In this book, we have not focused on ge*neral artificial intelligence*, which is defined as the hypothetical capacity of a machine to learn and reason about any cognitive task, as well as or better than a human. This hypothetical capability can be contrasted with the domain-specific capabilities of current AI systems. These current systems can acquire remarkable skills at playing two-person games or six-person Texas Hold'em, accurately predicting the weather or modeling the shape of a molecule. An algorithm designed to win poker against humans would not likely be able to predict the weather. Each algorithm is tuned to the parameters of its "game."

Discussions about artificial intelligence and its limits often lead to discussions about the Turing test. The Turing test is the iconic test of a machine's intelligence and, in particular, its conversational ability. The test is typically constructed to be "game" in which an AI software contestant attempts to be indistinguishable from a human. As illustrated in Figure 8-1, during the test an interrogator communicates with a machine and a human through text. No one can see the other, and the interrogator must decide who is the human. The interrogator sends a text message and the machine and human each send their separate replies. From the point of view of the machine and the human, the conversation is dyadic—they only know about their dialogue with the interrogator.

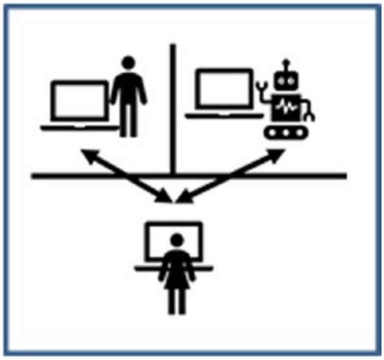

Figure 8-1. *A diagram of the classic Turing test*

As we write this, in March 2020, one of the authors asked Amazon's Alexa, "Alexa, can you talk to more than one person at a time?" Alexa answered, "Sorry I don't know that one." This was followed by, "Alexa, can you pass the Turing test?", to which Alexa replied, "I don't need to. I am not pretending to be a human." Conversational interfaces currently have limited ability to track the conversational flow in complex conversations and they typically cannot recall or make of use of prior conversations. Clearly, the version of Alexa that we accessed cannot pass the Turing test.

The developers of several conversational interfaces have claimed that their software has passed the Turing test, arguing that the Turing test is passed if a computer is mistaken for a human more than 30% of the time.

On June 7, 2014, *Eugene Goostman*, a software program that simulates a 13-year-old Ukrainian boy, was said to have passed the Turing test, a University of Reading competition.[10] On May 9, 2018, Google's CEO declared in reference to *Duplex*, Google's conversational voice technology, "In the domain of making appointments, it passes the Turing test."[11] The premier demonstration of Duplex was very impressive—it paused before responding, elongated certain vowels as if it were thinking, and inserted "uh" and "um," when appropriate.

Did these conversational interfaces pass the Turing test? We don't think so. As suggested by Harnad in 1992, the Turing test was not intended as a 5-minute game that can be won through clever distractions. The likely intent that was expressed through three variations of the "Imitation Game" was not to propose a 5-minute test of the ability to mimic human reasoning, conversation, or some other form of performance. The intent suggested by Harnad was that the Imitation Game was a thought experiment to demonstrate that the attribution of intelligence (human or otherwise) is not based on any deep intrinsic knowledge of other minds that is available after a short interaction, but is rather built up over many experiences. We cannot read minds, we can only judge behavior.

We bring this up, at the conclusion of this book, for three important reasons:

Firstly, the ability to mimic humans to confuse a judge about who is human and who is machine should not be the goal of collaborative robots or automation. Attempting to fool a human associate might be a serious ethical violation—it should always be clear when a decision or action is solely based on an algorithm; whether its investment advice, the reporting of a newsworthy event, or the far more serious judgment about someone's

[10]*BBC News* (June 9, 2014). Computer AI passes Turing test in "world first." www.bbc.com/news/technology-27762088 [accessed on March 25, 2020].

[11]Richard Nieva (May 10, 2018). Alphabet chairman says Google Duplex passes Turing test in one specific way. *CNET.* www.cnet.com/news/alphabet-chairman-says-google-duplex-passes-turing-test-in-one-specific-way-io-2018

innocence or guilt, it should always be clear to those that are impacted that the decision or action was the result of machine-based decisions.

Secondly, organizations and institutions often err in thinking that a machine intelligence would make better decisions or more objective, less biased decisions. Machine-based decision-making works best when the rules of the game are clear, as in a machine-learning system that plays chess or Go, or as in RPA where a business process is well defined, and each decision point has been considered by the process architect. The immediate danger of AI is not general superintelligence, but that institutions and businesses are “outsourcing” important decisions to machine-learning systems that are biased and limited by the data they process and by the domain-specific, single-purpose algorithms that drive their decision-making.

Thirdly, the algorithms that are hyped because they pass the Turing test often fail on closer inspection. As we worked on this book, it became clear that perceived progress in the domains discussed in the book has been greater than actual progress. This is supported by the research conducted by our colleagues in projects, other researchers around the world, and our own research. According to the media, autonomous vehicles are only a few years away, smart buildings are being constructed at a great rate, business processes are being supported by automation, and we will soon see customer facing and frontline operatives being completely replaced by conversational software robots.

These so-called advances also include medical robots that can replace doctors, robots managing end-to-end supply chains, and robot pickers in agriculture. A parent that one of the authors met described how scared she is for her 5-year-old daughter’s future because of all the jobs being lost to robots and automation. Her anxiety was easy to see and appreciate and in part provided some motivation for the book.

Working with Automation and Robots

Work in the future will change for many people and some areas of employment will be radically different over time. We see work being immediately affected by fewer jobs in transport, supply chain, and clerical tasks. This pressure is starting to be felt and will only increase.

Job losses are already being felt in repeatable clerical tasks and the use of tools like RPA is only going to accelerate this trend. One of the reasons for this is that RPA has a low cost of entry that is attractive to small- and medium-sized businesses. Training costs are comparatively low. A process can be automated more easily than a programmer can write a script. RPA is also attractive because of its ability to repeat a process in the same way every time without getting tired or bored and without making a mistake. Looking at literature online, it is clear that RPA has moved out of the lab, through testing, and is now an increasingly mature solution that is being sold to support business processes. In the RPA chapter we also discussed the strategy of keeping the solution either in the IT department or separate from the IT department and this can also have an effect on staffing levels. The pace of change has been accelerated by the COVID-19 pandemic and the work-only-from-home restrictions—hospitals, food distributors, and manufacturers are more willing to start pilot projects that introduce robots and automation.

However, there will also be new jobs created as people are released from menial tasks and are allowed more creative and sophisticated work. Work practices will also change, with a change in the balance of home working and commuting to new purpose-built smart buildings.

Artificial intelligence, machine learning, and deep learning are tools that can develop automation to the next level. Currently RPA can only execute existing processes. According to Serge Mankovski who was interviewed for the RPA chapter, AI tools will come into their own when process automation moves to process optimization. Intelligent automation should be able to examine both the business processes and their supporting infrastructure and optimize the whole process from end to end. This would cause disruption to the workforce since it would be able to take decisions in a more flexible way.

Successfully integrating robotic systems into the workplace requires careful examination of the goals and attitudes of those impacted by the new processes. Altering a business process may create efficiency in one area of concern but might create other logistical and social problems. When workers work alongside collaborative robots, they must be convinced that cobot is not recording every action and utterance or, if so, that the data will be kept private unless there is an extraordinary and compelling legal reason to analyze and expose it. This applies not only to social cobots that move and directly interact with humans but also to autonomous vehicles and smart buildings, and to robotic software that sifts through emails. Transparency and commitment to ethics is essential to a healthy work environment.

Creating a successful application takes time, patience, and money. It does not happen at "Internet speeds." A good example is self-driving cars and trucks. The expectation that they will be driverless and on the roads in just a year or so has been replaced by the understanding that change takes time. Due to regulation and technical difficulties, they will need a supervisor/driver for some years to come. An entire ecosystem of laws and regulations, of road-service providers, and of containers that are easily managed by mechanical hands needs to be created alongside the machine-learning algorithms.

In this new ecosystem, jobs may be lost, for example, in delivery and supply chain after automation, and robots are fully integrated. The job losses in these domains are frequently offset by productivity boosts and new work opportunities. When driven vehicles are replaced with driverless vehicles, opportunities will open up for more sophisticated servicing and maintenance. Computer engineers will be in demand to fix problems with the technology of the driverless car as well as mechanical engineers to fix the engine and brakes. Retraining will be an important factor in preparing for a driverless future.

Commuters in the future will have the chance to change their working practices. If you can work in a driverless vehicle, not being affected by motion sickness, you can leave your smart office earlier than normal and work all the way home. Suburbs will be pushed farther out as a long commute is the equivalent of working times. Changing these working practices may well result in better work/life balances and fewer stress-related illnesses as well as a reduction in road rage.

Employee health is a concern in some organizations that have a high rate of sickness absence. Smart buildings can provide personalized environments for workers, reducing the incidence of sick building syndrome. Smart buildings will also be able to remove the stress of parking at work by transmitting parking information to staff who are in driverless vehicles and on their way to work, assuming that the staff need parking information.

Data fusion is one of the tools that will allow all of this technology to deliver an integrated view of the work environment that can be understood by all stakeholders. Progress in high-quality machine learning and real-world models is still in the laboratories, but there are many research organizations working on this problem. The question of how the mix of digital, video, audio, radar, and GPS data can be gathered and fused into a single view is complex and the presentation of data fusion results will also be a challenge. Some of these solutions will take time to develop and commercialize, and the effects will not be felt in the next few years but over a much longer term.

Collaborative robots have the potential to increase the effectiveness of their human collaborators. Awareness of the autonomous entities in a collaborative team by its members will enable collaboration at a distance as well as in the immediate vicinity. For example, a collaborative team of robot and human bartenders and servers in a small space will use the same technology as a search-and-rescue team working over large distances. Simple robots such as robot vacuum cleaners will evolve to collaborate with human occupants, kitchen appliances, and waste disposal robots to manage a home, an office, or a factory. There are still complex technology problems to solve in maintaining safety and exercising judgment in decision-making. Many of these problems are, again, in the more distant future, but they will be solved. Society will evolve as collaborative robots evolve. In the future we will face as many societal challenges as there are technical challenges.

Final Thought

At talks about machine learning and robotics, we are often asked, “What advice can you give to those who are entering the workforce?” The answer is that work has been transformed by computers, by the Internet, and now by automation and tomorrow by AI-based robots. This does not mean that humans should compete with machines, rather we should be more human—the skills that are needed more than ever by industry from humans are curiosity, sociability, adaptability in thought and perspective, creativity, ethical judgment, and natural intelligence.

Index

A

P. Matthews and S. Greenspan, *Automation and Collaborative Robotics*,
https://doi.org/10.1007/978-1-4842-5964-1

B

C

D

E

S

T

U

V

W, X, Y

Z