Shaping Robotics Policy for the 21st Century

Insights from the 2016-17 Halcyon Dialogues

halcyon DIALOGUE

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

Foreword

Transformative technological shifts, such as the personal computer revolution and the rise of the Internet as a ubiquitous global communications medium, are almost always accompanied by unforeseen challenges and opportunities. As a result of the exponential growth in information technology, for example, entire industries now exist to safeguard critical infrastructure whose very existence was confined to the realm of science fiction just a short time ago.

Today, the field of robotics appears poised to undergo a similar episode of explosive growth, with intelligent autonomous systems promising to fundamentally alter the way humans interact with everything from transportation to medicine to the police. These developments have the potential to deliver great benefits to humanity, but to do so, they will have to be designed, tested, and employed in a manner consistent with an array of social, cultural, and economic parameters, many of which have yet to be fully articulated.

In light of the significant uncertainties associated with future developments in this emerging field, is it possible to create non-partisan policies that can help foster the development and adoption of robotic systems in a way that maximizes their advantages while mitigating the risks? We believe the answer is yes. The Halcyon Dialogue was conceived as a way to bring together global leaders, experts and researchers in a series of frank and open discussions on the challenges associated with robots and their broad implications for global society.

Over the course of four sessions in 2016 and 2017, an array of policymakers, innovators, and representatives from industry and academia convened at Halcyon, a creative space for 21st century problem-solvers in Georgetown, Washington DC. There, they examined the technical achievements of the evolving field of robotics; explored how the field is, or may be, affected by public opinion and policy; and discussed the broader social implications associated with integrating robots into the everyday lives of people.

The results of the dialogue have been compiled into this report, which was designed specifically with policymakers in mind. It is our hope that within its pages you will gain unbiased insight into the ways leaders in the field view the current state of the art in robotics from a diverse array of viewpoints, and that these in turn can inform sound policy.

AAAS's motto is "Advancing Science, Serving Society," and advances in robotics have the potential to profoundly influence the interplay between both the scientific and societal components of this mission in the years to come. AAAS and Halcyon are thrilled to have come together as partners on this endeavor. It is our profound hope that through this dialogue, we can help ensure that developments in this exciting and dynamic field take place in a way that advances human dignity and mitigates, rather than exacerbates, tensions in society.

Dr. Rush Holt, CEO, AAAS Dr. Sachiko Kuno, Founding Chair, Halcyon

Recommendations **Summary**

DATA, STANDARDS AND BEST PRACTICE

Define levels of autonomy in different domains and different contexts, as has been done with autonomous vehicles.

Establish better, more robust standards for the evaluation of device-specific training programs, including both human and technical factors.

Develop and incentivize the use of best practices for design and implementation and use third party authentication groups to enhance reliability and thus trust in robotic systems.

Standardize data collection and curation where possible, including on failures and near-misses, which would be used to develop best practices and communicate them to stakeholders.

Develop and implement best practices and a familiar design library that can be leveraged to maintain trust throughout the robotics and AI industries.

FUNDING AND INVESTMENT

Fund more robotics specialists in key safety research and regulatory environments to ensure that government agencies and industry have the expertise to anticipate safety problems.

Invest in organizations and partnerships that can bridge the gap between innovation and commercialization of robotics technologies.

Support research to examine the effects of robotics on workforce, human behavior, the economy and society as a whole.

Increase fellowships in government to connect scientific and technical experts with policy makers and vice versa.

EDUCATION AND OUTREACH

Create an Advanced Research Projects Agency - Education (ARPA-ED) to explore innovative and technology-intensive approaches to robotics in education at the K-12, undergraduate, and graduate levels.

Pursue public education programs to foster widespread understanding of the capabilities and limitations of robotic systems.

Apply the methods and tools associated with robotics, such as machine learning, augmented reality, and virtual reality, to radically improve the quality of education.

Institute modern versions of the industrial arts in the K-12 curriculum as robotics leads to more mechanized workplaces.

OPPORTUNITIES FOR ADDITIONAL DIALOGUE

Foster interactions among the users, developers, and producers of robotic systems, including public education, so that all parties understand the functions and limitations of these new technologies.

Convene a group of high-level leaders to survey the field of robotics with both breadth and depth and communicate through campaigns grounded in productive discussions and visualizations.

Encourage international exchanges of information among countries to reduce duplication and accelerate innovation.

Incorporate a diversity of voices in testing regimes, so that people of all backgrounds, ages, and experiences are heard and factored into the decision-making process.

RESEARCH AND STUDY

Develop ways for robotic systems to communicate the confidence they have in making a decision before action is initiated.

Study the ways in which robotic systems can fail, since the increasing complexity of systems can make them fail in unanticipated ways.

Determine whether incremental improvements to a robotic system can be tested just for safety and not for both efficacy and safety.

Explore the extent to which some degree of simulation can act as a proxy for other forms of regulatory scrutiny.

Make use of national laboratories in transferring robotic technologies from the military to the public sector.

Design robots so that communication to implement their behaviors comes through multi-sensory signals such as sounds, lights, and motions, similar to the way that car brake lights are used to indicate a reduction in speed.

GOVERNANCE AND REGULATORY

Require transparency for testing regimes so that the public is informed about testing practices.

Look to other communities, such as the medical and automotive sectors, for lessons in developing regulatory systems.

Develop and communicate reliable maintenance procedures for robotic devices to enhance safety.

Create a stratified access system analogous to the military security system with different levels of protection.

Address fundamental ethical and regulatory questions earlier in the process of transitioning military robotic systems to the civilian sector.

Use a variety of policy tools to support and encourage the responsible development and application of robotic systems, such as tax incentives, support of research and development, and purchasing subsidies.

Work with stakeholders, including the FDA, to determine the appropriate regulation of robotics.

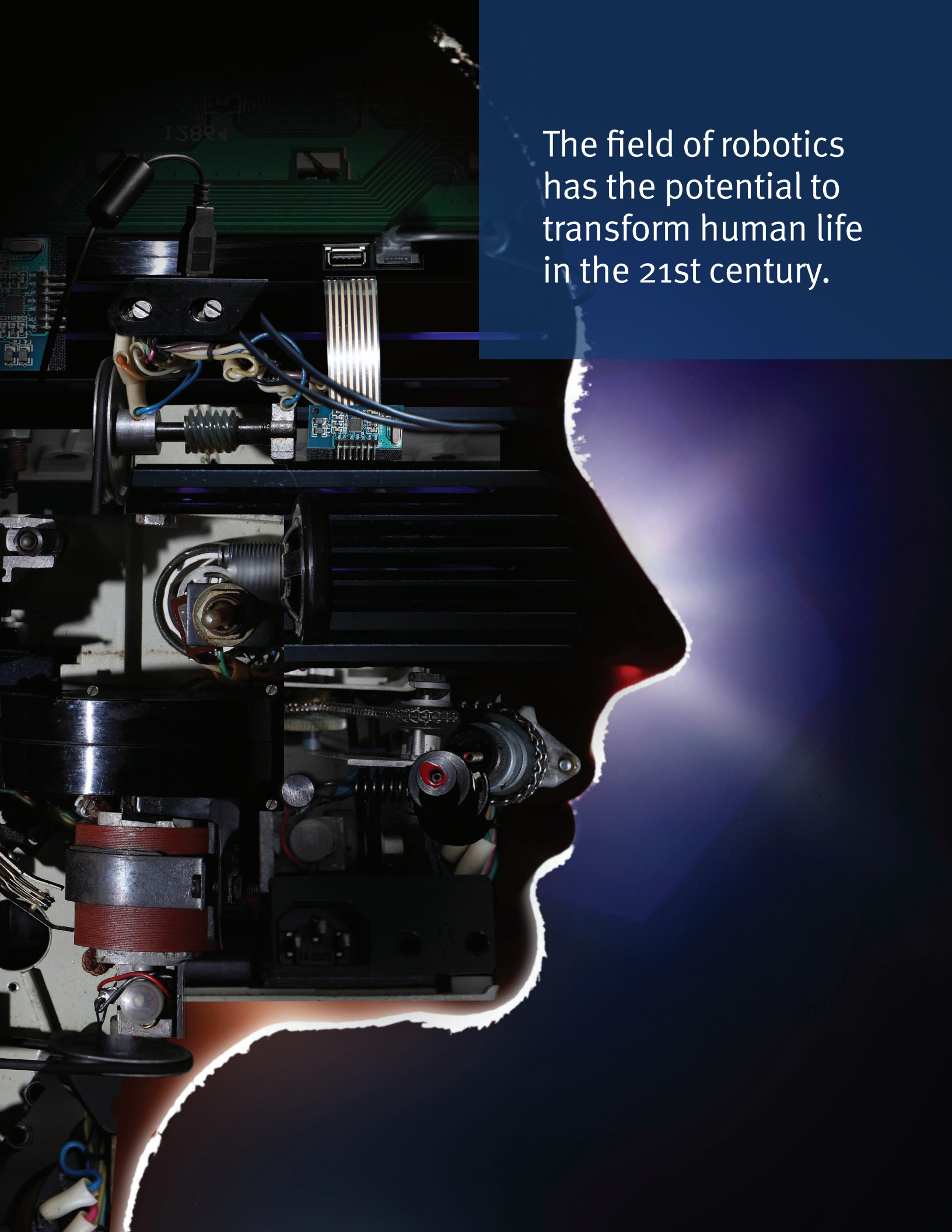
Make FDA filings more open so that judgments about efficacy and safety can be more widely agreed upon and accepted by companies and the public.

Disclose to patients of robotic medicine the levels of autonomy, future benefits, and potential risks involved in a particular medical procedure and discuss in advance the alternatives to that procedure.

Establish new legal and regulatory provisions to ensure that robotic and AI data are not misused.

Create a federal interagency coordinating body to further the development of legal, regulatory, and ethical decision making.

In situations where producers of robotic systems self-regulate, determine the implications and enforcement mechanisms for failing to adhere to those regulations.



The field of robotics has the potential to transform human life in the 21st century.

Executive **Summary**

As mechanical engineering, material science, electronic engineering, bioengineering, information technologies, artificial intelligence, and other contributing fields continue to advance, robotic technologies could become increasingly common in workplaces, homes, and public spaces. The resulting societal changes could be as dramatic as those seen in the 20th century.[1]

Given the enormous potential of robotics to transform society, Halcyon and the American Association for the Advancement of Science held a series of four high-level discussions among global leaders, experts, researchers, and innovators from October 2016 through June 2017 to identify the issues that are fundamental to robotics and to propose actions that could shape the field and its societal consequences. The discussions were held under the Chatham House Rule, which permits the free use of the information provided at a meeting but without attribution to the individuals providing that information. From those four meetings, nine major themes, with corresponding recommended actions for policy makers, academic researchers, companies, and nonprofit organizations, emerged.

Determining Effects on the Workforce

Robots could have a dramatic impact on labor markets. To date, the effects of robots have been most visible in manufacturing, but robots eventually could do many jobs currently being done by people in other sectors of the economy, including health care and service sectors.

Technological advances in the field of robotics and the subsequent adoption of robotic technologies in the private sector will change the nature, numbers, and structure of jobs throughout the economy, though exactly how negative and positive effects will be distributed remains highly uncertain. Robots will expand many job categories and create entirely new kinds of jobs; by increasing productivity, they could allow industries to expand and compete globally. They also could

RECOMMENDATIONS

Support *research to examine the effects of robotics on the workforce and on the broader society. Fundamental public policy questions include: What is the pace of change? Who is being affected by change? How are they being affected? Where are changes happening?*

1 The documents listed in the "Additional Reading" section at the end of this report describe some of the future changes that can be expected in the age of robotics.

reduce demand for some categories of jobs, along with the jobs of ancillary workers who today support those economic sectors.

The effects of automation will vary greatly by sector, by region, and by country. Robots could be competitive with low-skilled, repetitive jobs overseas. On the other hand, countries with less experience with robotics could leapfrog more advanced countries, especially if their educational systems build expertise, a skill base, and enthusiasm for robotics.

A major question is whether robotics will increase or decrease inequities among groups of people or nations. For example, the returns on investments in robotics could flow to a relatively narrow portion of the population. Greater use of robots could even alter social structures, such as gender roles, as they change what people do in workplaces and at home.

Enhancing Design and Implementation

The design and implementation of robots are critical factors in how they will be used and whether their use will be accepted or resisted. For example, if robots are designed and presented as tools, people will be less likely to be afraid of them or see them as threatening.

The social expressivity of robots will be an important factor in their use and acceptance. A robot that is aware of the social, cultural, and gender cues of users can be more responsive and will be more readily accepted. On the negative side, if robots are perceived as social entities, people could divulge private information to robots, or they could be deceived by robots.

Design and implementation also have a direct impact on safety and, therefore, on the trust people have in robots. A robotic device may be safer than one operated by a human, but no device is error free. Both designers and operators therefore need to think about what will happen if something goes wrong. In some cases, legal and regulatory frameworks will be needed to enhance safety and maintain trust.

The users of robots are the experts in that use, not the designers of robots. By engaging with users, designers can take advantage of that expertise. Involving the users in design reveals what a person or community needs, which provides a way to build value and sustainability into a robotic system.

RECOMMENDATIONS

Foster *continued dialogues among people representing different points of view — including, in particular, users — to create more representative design processes.*

Develop *and incentivize the use of best practices for design and implementation and use third party authentication groups to enhance reliability and thus trust in robotic systems.*

Building Trust

To earn trust, the public will have to perceive that the performance of an autonomous robot is better than human performance in terms of reliability, efficiency, and safety. This will require extensive testing, both through simulations and in the real world. Trust may also require the enactment of laws and regulations, which then have to be enforced in a fair and reasonable manner.

The extent to which the public is informed, involved, and engaged will determine the level of transparency with robotic systems and, in turn, will influence public trust. Comfort levels with autonomous systems are different for different populations. Adults are more familiar with robots now than they used to be because their children or grandchildren are learning robotics in school. Investments in education and training for all demographic groups may be as crucial as investments in technology in engendering trust.

Trust in robotic technologies will initially and primarily be built through the teaming of people with autonomous systems. Trust in robots will grow gradually and through small steps, with small and repeated transactions leading to more significant ones.

RECOMMENDATIONS

Foster *interactions among the users, developers, and producers of robotic systems, including public education, so that all parties understand the functions and limitations of these new technologies.*

Introduce *new systems deliberately and with options for implementation that make the transition more manageable.*

Minimizing Risks

Minimizing the risks of robotic systems requires testing and evaluating those systems. Currently, hardware is tested in a comprehensive and deterministic manner, but exhaustively testing the entire realm of possibilities in which an autonomous system can operate will often be impossible. Determining the borders within which robotic systems have the potential to fail will require trial and error, accidents, and incidents that will test the public's trust in these systems.

Engendering trust in robotic systems will be important to their integration into society, but overly trusting these systems also will create problems. Trusting robotic systems to perform beyond their capabilities or attributing personalities and relationships to these systems may encourage improper use. People have already made this mistake by overly trusting autopilots in semi-autonomous vehicles despite access to information on the vehicles' limitations.

Transparency in testing will encourage public trust and the adoption of new systems. A feedback loop among affected communities, developers, and other stakeholders will provide a more effective and supported integration of these systems into society.

RECOMMENDATIONS

Require *transparency for testing regimes so that the public is informed about testing practices.*

Incorporate *a diversity of voices in testing regimes, so that people of all backgrounds, ages, and experiences are heard and factored into the decision-making process.*

Defining Autonomy

Autonomy in robotic systems is not a binary condition. It is a spectrum of capabilities and depends on the context in which robots are used. For example, robots can be more autonomous in some areas of health care, such as nursing or rehabilitation, than in areas such as surgery. Robots also can support humans in making decisions, creating what might be called supervisory autonomy.

Autonomy can depend as much on a user's perceptions as on a technical definition. For example, when a user is controlling a robot with a joystick, the robot may be performing calculations to respond to commands, even though this level of autonomy may not be visible to the user.

RECOMMENDATIONS

Define *levels of autonomy in different domains and different contexts within those domains, as has been done with autonomous vehicles.*

Develop *ways for robotic systems to communicate the confidence they have in making a decision before action is initiated.*

Issues associated with safety and risk arise with robotic autonomy. A robotic system may not be transparent, and it may not be possible to determine exactly what has happened if something goes wrong. A robot cannot participate in a failure analysis in the same way that a human can. People can understand another person's thinking processes in ways that they may not be able to understand a robot's processes.

Collecting and Analyzing Data

Robotic systems can collect vast amounts of information. These data could be instrumental in driving leaps in robotic capability through machine learning algorithms or other techniques powered by large data sets. However, in some use cases, technical, administrative, or ethical challenges could stand in the way of this vision.

One way to overcome these challenges is through the development of standards for the collection, tagging, expression, storage, and integration of data. Certification requirements then could govern the collection of particular types of data. Data also need to be transparent, so that broadly based judgments can be made based on readily available data.

RECOMMENDATIONS

Standardize *data collection and curation where possible.*

Incentivize *collection and reporting of data on failures and near misses.*

The airline industry has created mechanisms for anonymously reporting near misses, which gives industry and regulators better data with which to analyze flaws and improve systems. Robotics could help flag near misses after the fact to gain more information about safety.

Supporting Education and Training

The age of robotics will accelerate and intensify many of the trends in education and training initiated by the information age. In the future, the majority of jobs could involve some aspects of engineering and computer science, and education will need to move in that direction.

As the economy changes, people will need scientific and technical skills to fill many jobs, but they also will need what have been called soft skills, such as the ability to learn quickly, to communicate effectively, and to work in teams. With robotics, for instance, combining engineering with the arts and social sciences could lead to designs for service robots that address not just physical needs but emotional needs, resulting in robots with which people feel more comfortable and receptive.

An option being pursued by governments in some other countries is to subsidize the expenses an individual incurs each year on retraining, which emphasizes the message that people are responsible for their own retraining and need to think continually about how to update their skills. However, retraining is not an option for everyone, and in some places the term itself is disparaged because of unsatisfactory results in the past. Some industries and communities will continue to lose jobs, and not everyone in those industries or communities will learn the new skills that will be needed or adapt to new workplaces. Better public policies will be needed to deal with the inevitable human consequences of roboticization.

RECOMMENDATIONS

Apply *the methods and tools associated with robotics, such as machine learning, augmented reality, and virtual reality, to radically improve the quality of education.*

Institute *modern versions of the industrial arts in the K-12 curriculum as robotics leads to more mechanized workplaces.*

Ensuring Effective Governance

Achieving public trust and legislative action without stifling innovation and competition is a difficult task. New regulations may be needed; some existing regulations may need to be changed or eliminated. Governments could encourage the producers of robotic systems to self-regulate; they could enable a variety of approaches and then develop some kind of performance evaluation framework to determine whether a device meets safety standards; or they could rely on private standards development organizations, hundreds of which form and enforce industry standards.

One complication is that the issues are inherently international. Robots can be operated remotely from anywhere in the world. How can national laws and regulations be respected in an inherently international setting?

Public policies also can encourage the growth and development of new technologies through such measures as procurement decisions, tax policy, liability limitations, and support for research. Governments can guide the development of a technology without creating a rigorous framework for its future. For example, governments can support studies that analyze possible pathways into the future and how to influence future developments. Increasing science diplomacy between countries can reduce uncertainty on an international level.

RECOMMENDATIONS

Use *a variety of policy tools to support and encourage the responsible development and application of robotic systems, such as tax incentives, support of research and development, and purchasing subsidies.*

Encourage *multi-stakeholder dialogues along with conversations arranged for particular purposes, such as establishing industry standards or best practices.*

Creating Advisory Structures

A much larger breadth of engagement is needed to encompass the range of people who will be affected by robotics. Without feedback from across society, manufacturers will make products that may not work for a substantial portion of the population. Increased engagement will allow for a system of accountability, input, and improvement.

Nearly every sector of society should have a voice in considering the oversight of robotic technologies, since all sectors and individuals will be affected. If members of the public perceive that only one part of society has input into regulations, they will distrust the results.

More broadly distributed and deeper expertise can be developed through fellowship programs and science diplomacy. Academia, government, and private companies all have roles to play in maintaining the United States' leadership in the development of robotics.

RECOMMENDATIONS

Create *an independent organization, or augment an existing organization, to provide assistance to the robotics sector in managing risks and encouraging best practices.*

Convene *a group of high-level leaders to survey the field of robotics and convey information about robotics grounded in continued productive discussions.*


The rapidly expanding use of automation, artificial intelligence, and robotics could reshape human life . . .

1. Introduction

in the 21st century. Robots are already used extensively in manufacturing, packing, transportation, Earth and space exploration, health care, weaponry, laboratory research, safety, and many other fields. In the future, their applications will be essentially limitless.

To explore the future of robotics and identify actions that can be taken now to prepare for that future, the public nonprofit organization Halcyon, through its Halcyon Dialogue Program, and the American Association for the Advancement of Science (AAAS) held a series of four meetings from October 2016 through June 2017 at Halcyon's headquarters, Halcyon House, in the Georgetown neighborhood of Washington, DC.[2] Each of these Halcyon Dialogues on Robotics brought together approximately 20 global leaders, experts, and innovators in the field of robotics, representing a broad range of sectors, countries, backgrounds, and ages, to discuss the implications of robotics for global society in four broad areas:

- The promise and peril of military robotics technology in civilian settings
- Evolving capabilities and impact of robots in medicine
- Global ramifications of robotics for work and social justice
- Emerging legal, legislative, and liability issues at the intersection of robotics and policy

The Appendix lists the Halycon Dialogue Sessions meeting participants and the program committee members, who identified the issues to be discussed, suggested how the discussion should be formatted, and recommended participants. Each meeting was led by a moderator, with the participants identifying the major issues that emerged at each meeting (as reflected in the headings of the chapters in this report). For the purposes of the meetings, the programming committee defined robots as "autonomous or semi-autonomous systems that interact directly with the physical world." Per this definition, certain elements of the "internet of things" were included, but "bots" consisting purely of software were not considered. This focus, however, did not exclude consideration of software as essential to the functioning of robots.

---

2 The Halcyon Dialogues on Robotics were funded by Mitsubishi Heavy Industries, Ltd.; Honeywell; Hitachi, Ltd.; Microsoft.; and X, the moonshot factory. The observations, conclusions, and recommendations contained in this report are those expressed by participants at the four meetings and do not necessarily reflect the views of the funders, Halcyon, or the AAAS, its Board of Directors, its Council, or its membership.

This report summarizes the discussions of the four meetings, with the addition of background material on the sectors examined in the first two meetings. Chapters 2 and 3 consider the impact of robotics in health care and the civilian use of military robotic systems, respectively, as a way of introducing the broader issues associated with robotics in general. These chapters address such questions as: How much autonomy should medical robots have?[3] What are the appropriate ways for government to use military-derived platforms, both lethal and non-lethal, in domestic settings? What are appropriate safeguards to protect the public from misuse?

> Robots are not inherently good or bad. However, they can have effects that are positive or negative for particular people and groups of people.

3 This question was also addressed following the second dialogue in an editorial published in Science Robotics by a number of dialogue participants (Science Robotics 15 Mar 2017: Vol. 2, Issue 4, eaam8638 DOI: 10.1126/scirobotics.aam8638)

Chapter 4 looks at several of the broader issues raised in the first two meetings outside of specific contexts. It reflects on automation and the workforce throughout the economy and in the United States and other countries, the design and implementation of robotics for the greater good, and the broader impacts of robotics on society.

Chapter 5 turns specifically to the legal and regulatory issues associated with robotics. It considers features designed to minimize risks and gather data from robotic systems. It also examines the role of government in establishing standards, laws, regulations, and ways of providing input to the policy-making process.

Chapters 2 through 5 conclude with recommended actions developed during small-group sessions and proposed during plenary discussions at the meetings. These recommended actions should not be seen as consensus recommendations from the meeting participants since the meetings were not structured to produce such a consensus. Rather, they should be viewed as ideas emerging from the meetings that deserve to be pursued, both through further discussion and through implementation.

Robotics is a contentious topic, but the areas of controversy do not necessarily focus on the most important issues. Robots are not inherently good or bad. However, they can have effects that are positive or negative for particular people and groups of people. The Halcyon Dialogues on Robotics were designed to identify the issues that are fundamental to robotics and to propose actions that can shape the future of robotics and its consequences throughout society.


With health care costs accounting for one-sixth of the U.S. economy, improving human health . . .

and well-being has been and will continue to be a major focus of robotics. Already, robots serve in a wide variety of areas, including surgery, home care, rehabilitation, prosthetics, and hospital automation. These and other application areas are likely to see rapid growth.

This chapter, which is based on the second Halcyon Dialogue on Robotics, reviews some of the uses of robots in health care and then examines three context-specific issues: safety and regulation, autonomy, and the collection and use of data.

The Capabilities of Robots in Health Care

Of the many applications of robots in medicine, surgery has drawn the most attention. Computer-assisted surgical systems have become mainstream treatment options in recent decades as clinicians recognize the distinct advantages that these methods provide. Surgical robotic devices allow surgeons to use smaller instruments, minimize incision size, and reduce pain, blood loss, the risk of infection, scars, and recovery time. With the da Vinci Surgical System, for example, surgeons have control over four slave arm manipulators, three of which are surgical tools and one that is an endoscopic camera. The tiny wristed instruments enable the surgeon to operate with a greater range of motion than the human hand.

As the use of robots broadens to include more advanced procedures, difficult issues are likely to arise. For example, multiple parties are involved in the end-to-end operation of a robot, including the manufacturer, the surgical team, the hospital, the maintenance provider, and the software developer, which can complicate determinations of liability. One proposed option is to use a black box system like those used on aircraft that would record all movements, commands, and executions involved in a robot-assisted operation, without possible modification or manipulation from the outside. Robotic systems could also monitor their own operations and save the information for later analysis. This information could be used later to investigate anything that has gone wrong and to improve safety, just as that information has been used in aviation to make flying safer. This also would help build the trust that people will need to adopt robotics.

HOME CARE AND REHABILITATION

Beyond surgery, a huge gap exists in home health care that could be filled with semi-autonomous robots. Care and rehabilitation robots could deliver consistent, lengthy, and personalized therapy without tiring; conduct exercises not possible by a human therapist; implement continuous, adaptable, and focused treatment plans; or otherwise augment conventional approaches. They could assist in movement disorders, such as those resulting from stroke, traumatic brain injury, or other trauma. They could also act as intervention and therapeutic tools for social and behavioral disorders, including autism spectrum disorder, attention deficit hyperactivity disorder, and addiction.

Robots could increase quality of life and support life-long independence. This includes improving mobility, reducing isolation and depression, and enhancing the ability to age in place. Care robots could clean patients, fetch items, dispense medications, serve as memory aids, give physical assistance, and provide psychological support. Robotic fulfillment of a wide variety of tasks would free up human caregivers to more properly address the psychological needs of those in their care. Already, care robots are being used in homes for the elderly. However, many of the tasks involved in elder care are very complex, such as dealing with someone who has Alzheimer's disease or a serious disability. Much of this care today takes place largely out of sight in facilities or in homes. Perhaps robots will be able to assume some portion of those responsibilities, but probably not anytime soon. In that case, the most pressing issue is how to get proper care for individuals who need it now. Elderly patients may be able to live independently with robotic assistance, but human support may have psychological advantages that robots are unable to provide.

In the area of rehabilitation, robots could allow people to exert forces in more complex ways and for longer durations than human therapists can. At a higher level, robots could evaluate a person's performance from session to session and generate feedback about what is effective and what is not effective. In this way, therapy sessions could gradually become more autonomous, with an increasing number of interactions occurring between patients and robots.

SOCIAL INTERACTIONS

A new and developing model in care and rehabilitation is the socially assistive robot that focuses on using sensory data about the user to determine appropriate behavior. Data compiled from a multitude of sensors on the robot, in the environment, and worn by the user can be incorporated into statistical methods for user modeling. These robots are designed to assist users through social interactions, as well as or rather than, physical interactions. This technology could enhance the quality of life for very large populations, including older adults, people with cognitive impairments, people rehabilitating from disease or disability, and children with socio-developmental disorders. Socially interactive robots also could be used in the diagnosis of behavioral disorders, which often requires prolonged periods of observation. Long-term interactions

will require socially engaged robots to adapt their behavior to changes in a user's state of mind, in responsiveness to different behavioral strategies, and in the relationship established between the robot and its user.

Many people foresee social interaction as a key feature of the future of robotics. Human emotion is understood through a combination of factors, including voice, facial expression, body motion, gestures, and physiologic data. Models that capture the complex patterns of social behaviors and interactions will have to be designed, but once these models exist, the potential tasks that robots can accomplish will significantly increase. Robots will be able to act as consultants, therapists, friends, and caregivers. They will be able to express authority, compassion, or competition.

The emotional responses to robots intersect with safety and ethical concerns. When robots perform not only a therapeutic but a social function, this could be used to exploit some people, such as children or people with dementia. The ethical issues associated with forging an emotional connection with a machine are longstanding concerns in robotics and artificial intelligence.

OTHER APPLICATIONS

On the user side, brain-computer interfaces have been developed that allow operators to control machines solely through brain activity. Such systems are most commonly used to assist, augment, or replace human cognitive and motor functions, including the restoration of damaged hearing, sight, or movement through prosthetic limbs or remote control.

Many people foresee social interaction as a key feature of the future of robotics.

Robotic prostheses and orthoses can protect, support, or improve the function of various parts of the body. Robotic prostheses are artificial extensions of a person's body that replace an absent body part by fusing mechanical devices with human muscle, skeleton, and nervous systems. They eventually aim to emulate the missing body part through replication of the many joints and limb segments and seamless neural integration that provides intuitive control of the limb as well as touch feedback. Entire exoskeletons paired with brain-computer interfaces have been developed to allow paralyzed individuals to walk again.

As robots become part of the culture, they will increasingly be related to modes of expression and artistic preferences. Some people want prosthetic limbs that closely approximate human limbs; others want clear limbs or high-tech limbs. The same function can be served differently for different people or in different settings.

Micro nanorobotics is another emerging medical field in which robots built at a micro scale with nanometer features can carry out functions within the human body. In the future, these microscopic machines may be able to identify and destroy cancerous cells, deliver drugs, and carry out targeted surgery. They may also lead to advances in genomics, brain mapping, and other forms of data gathering from biological systems.

Telepresence has the potential to expand tremendously as it allows specialized doctors to diagnose and treat patients at a distance. Such systems could improve acute and postoperative care, provide long-term management of chronic conditions, and offer an alternative to residential living facilities for elderly patients. Telemedicine also could reduce the inequity gap by providing treatment for people who live outside populated areas, and it can be used in disaster contexts and in environments that are far from medical personnel.

Finally, robots are extensively used in medical education. Each year over 180,000 doctors, nurses, emergency medical technicians, and other first responders in the United States train on high-fidelity robotic patient simulators — life-sized mannequins that can breathe, bleed, respond to medication, and interact with learners. These simulators help clinicians practice procedural and communication skills before treating actual patients. Additionally, surgeons of all disciplines use task trainers — lifelike models of specific anatomical regions — to perfect their motor skills and learn new procedures.

LIMITATIONS AND QUESTIONS

Current robots have many limitations, and they pose difficult decisions in health care. Robotics brings a new level of complexity to some medical procedures. They raise questions about the credentialing of not only the devices but the users of those devices; for example, no standard requirements exist for acquiring credentials in robotic surgery after the completion of one's residency. Brain-computer interfaces and microrobots usually involve the introduction of technology into the body, creating a complex set of liability issues. Currently, the United States regulates nanotechnology by reference to size and delivery method. Thus, nanorobots acting through chemical means are defined as drugs while

those acting through physical means are defined as devices, a distinction that has legal and economic implications. Telepresence medicine also will create liability issues, since the jurisdiction will have to be determined for region-specific liability laws.

Other advances could fuel different kinds of controversies. So far, the regulation of medical robotics has focused largely on their use in identifying and treating medical conditions and replacing lost natural function. But technological developments may enable non-necessary medical advances that have ethical implications. For instance, what if an exoskeleton could allow a disabled person to run farther or faster than a normal human, or if a robotic hand could allow people to lift heavier weights or carry out tasks with greater accuracy than before? Biological enhancements that draw on robotics will raise issues beyond those associated with the treatment of disease or disability.

Safety and Regulation

Among the most important issues surrounding the use of robots in health care are safety and trust. (The latter subject is discussed more extensively in the next chapter.) Is the use of a particular robotic technology safe? Do people trust a technology to be safe?

A device used in health care may be safer than an intervention delivered by a human, but no device is error free. Human operators, therefore, need to think about what will happen if something goes wrong. Anticipating failure can be especially complicated with robotics. Usually, failures are not mechanical failures. Rather they are unanticipated consequences of the design of a robot, such as an algorithmic quirk. This relates to the issue of resilience and how an environment can be made resilient enough to allow for contingency.

Error rates are not just absolute but also are related to user expectations. Thus, many people think that they are much better drivers than they actually are. When told that an autonomous vehicle has a low error rate, they may still view the risk as unacceptable out of the mistaken belief that they have an error rate lower than the vehicle.

Simple fixes can promote safety, such as not allowing a robot to operate too close to a vital nerve or artery. However, people may then rely on the safety features of a robot rather than their own expertise to prevent them from making errors, which could increase the rate of unanticipated errors.

Risk analysis requires analysis of the alternatives, but formal education and training in risk analysis and mitigation is rare for both the designers and users of robots. Most people learn those skills on the job or directly from the regulatory agencies with which they work. Training in this area may be important for the greater introduction of robotics into society. Risk analysis also requires knowing the risks of current procedures, which will be an important area for discussion and investigation.

Recommendations Related to Safety and Regulation

Study *the ways in which robotic systems can fail, since the increasing complexity of systems can make them fail in unanticipated ways.*

Determine *who should be making judgments about efficacy and safety and how those judgments should be made.*

Incentivize *reporting and collection by professional societies of anonymized registries of failures and near misses. These data would be used only to develop best practices and communicate them to stakeholders. At present, these data are underreported despite their potential to improve safety.*

Find *ways to encourage at least temporary storage of exhaustive data profiles that can be accessed in case of failures or near misses. An analogy is a surveillance camera in a store that records data for 24 hours so that, if something happens, recent data are available to document the event.*

Develop *and communicate reliable maintenance procedures for robotic devices to enhance safety. This could occur through a federal agency, through industry consortia, or through public-private collaborations.*

Determine *whether specific recommendations can be eliminated without jeopardizing efficacy or safety.*

Work *with stakeholders, including the FDA, to determine the appropriate regulation of robotics.*

Make *FDA filings more open so that judgments about efficacy and safety can be more widely agreed upon and accepted by companies and the public.*

Determine *whether incremental improvements to a robotic system can be tested just for safety and not for both efficacy and safety. For example, if the hardware or software behind a system has been approved for both safety and efficacy before, is the same degree of critical scrutiny warranted, or can less stringent standards be applied?*

Explore *the extent to which some degree of simulation can act as a proxy for other forms of regulatory scrutiny.*

Fund *more robotics specialists in key safety research and regulatory environments to ensure that government agencies and industry have the expertise to anticipate safety problems and develop best practices.*

Establish *better, more robust standards for the evaluation of device-specific training programs, including both human and technical factors.*

In some cases, legal and regulatory frameworks will be necessary for the development and implementation of robots in medicine. (Chapter 5 discusses laws and regulations in more detail.) Such frameworks can create a clear risk and liability profile so that people can figure out how they want to use and invest in technology. For example, a local institutional review board can make it more difficult for a technology to be adopted, but it also can provide protection for the introduction of a technology and warn of possible problems to be anticipated. A broader regulatory framework could hasten rather than slow progress, especially if procedures can be made clearer or more streamlined.

Regulatory oversight can differ greatly from one institution and location to another, which makes coordination difficult. For example, institutional review board approvals that apply beyond a single institution would be a huge advance. In addition, regulatory agencies and government need people with expertise, which requires both previous experience and continued research to more fully understand safety and efficacy.

Professional medical societies are much more involved than government regulatory bodies in controlling the practice of medicine. For example, they develop clinical practice guidelines and registries of cases, which can provide a data set for types of devices and how they perform. Such guidelines and registries could apply to applications of robotics as well.

Autonomy

Autonomy in medical robotics is not a binary condition. It is a spectrum of capabilities. Autonomy also varies by context — for example, between home care robots and surgical robots. In general, autonomy remains a poorly defined term in medical robotics.

In some cases, the degree of autonomy ascribed to a robotic system depends as much on a user's perceptions of autonomy as on a technical definition. For example, when a user is controlling a robot with a joystick, the robot may be performing calculations to respond to commands, even though this level of autonomy may not be visible to the user. The robot is not making human-level judgments, but it is still acting autonomously within certain boundaries.

Autonomy is also an issue when a surgeon in an operating room decides to delegate part of a surgery to a resident or medical student. But distinct issues associated with safety and risk arise with robotic autonomy. A robotic system may not be transparent, and it may not be possible to interrogate that system if something goes wrong. A robot cannot participate in a failure analysis in the same way that a human can. People can understand another person's thinking processes in ways that they may not be able to understand a robot's processes.

This raises the issue of "contextual trust." Robots might be trusted in particular contexts but not in others. A robot can be very good at a specific task but would not be trusted to learn and perform other tasks. Relying on contextual trust can enable innovation to continue without excessive regulatory oversight.

Today, medical robots generally combine machine capability with human judgment, and that is unlikely to change in the near future. Machines can assist a surgeon in doing what the surgeon wants the machine to do, but decision making is unlikely to move away from the physician — at least in the short term. This is the case with current radiation therapy, where a computer is used to compute machine settings to deliver a radiation dose prescription written by the physician.

Although some health care activities can be isolated and turned over to autonomous systems, such environments tend to be very constrained. Machines can be more autonomous in some areas, such areas as care giving or nursing, but even in these cases many tasks require difficult cognitive assessments. Robots may also support humans in making decisions, in which case robots may have some degree of autonomy with human oversight.

The medical legal system will be an important influence on the development of autonomous systems in medicine. Companies' and physicians' fears of being sued can impede the development and implementation of systems. Machines may not be able to make decisions on their own for legal reasons, not technical reasons.

An important step will be to define levels of autonomy for applications of medical robotics. These levels may range from choosing the best way to implement a straightforward command to dealing with unforeseen conditions and circumstances.

Recommendations Related to Autonomy

Define *levels of autonomy in medical robotics, as has been done in transportation. Levels of autonomy will need to be defined differently in different contexts, and expertise will not necessarily be shared across domains.*

Disclose *to patients the levels of autonomy, future benefits, and potential risks involved in a particular procedure and discuss in advance the alternatives to that procedure.*

Develop *ways for robotic systems to communicate the confidence they have in making a decision before any action is taken.*

Collecting and Using Data

The final theme discussed during the meeting on robotics in health care involves data: data collection, data transparency, and big data. Medical data could be instrumental in driving leaps in robotic capability through machine learning algorithms or other techniques powered by large data sets. The use of robots would make it possible to know exactly what medical procedure was performed. That information can be correlated with outcomes to figure out the best ways to treat or prevent disease.

But, in some cases, technical, administrative, or ethical challenges could stand in the way of this vision. The data are not of uniform quality. They are not as rigorously controlled as in a clinical trial. Hospitals can be protective of their own data, whether for legal, commercial, or cultural reasons. Patient privacy needs to be protected, both through provisions of the Health Insurance Portability and Accountability Act (HIPAA) and through other privacy protections. Collecting high-quality data can be very expensive, which is why clinical trials are so expensive. Data need to be curated and controlled in significantly constrained environments. Even such a simple matter as expressing a date can be done very differently. Standardization can be difficult among institutions and across countries. Each can have its own procedures, its own review boards, its own lawyers, and its own ideas about intellectual property. For instance, Canadian physicians do not collect a patient's date of birth, because that is seen as a unique identifier, whereas U.S. physicians do so routinely. Such differences can make it difficult to gather and compile large amounts of data and make inferences from the data.

One way to address such problems is through the development of standards for the collection, tagging, expression, storage, and integration of data. Certification requirements then could govern the collection of particular types of data. Another approach is to embed data records within medical records so that data collected for other purposes can be extracted for research.

In the United States, the Food and Drug Administration (FDA) generally has input into the design of studies and the kind of data being collected, providing assurance that the quality of the data is good. For example, it has teams of people who audit the data being collected at clinical sites. The agency also inspects manufacturers that are collecting data and building devices, again to provide assurance that the manufacturing process is adhering to a certain level of quality. Even before a clinical trial, a device has to go through a quality review to demonstrate that the device exhibits a certain level of performance, reliability, and safety.

Medical robots will make possible the collection of huge quantities of data, and big data analyses could draw useful insights from the data. Like autonomy, big data is a term that encompasses many levels and kinds of information. In some cases, even a few hundred or thousand patients can provide very useful data — after all, surgeons are trained on far fewer patients. In other cases, much larger numbers may be necessary to draw useful conclusions.

Medical robots will make possible the collection of huge quantities of data.

The reliability of the conclusions drawn from the use of big data is an important issue. Even when data analysis points in a particular direction, the conclusions drawn from the data will need to be tested. The data will not always speak for themselves.

Data need to be not only available but transparent, so that broadly based judgments can be made about conclusions based on readily available data. Openly available data allow others to assess the quality of the data and decisions based on the data. In this way, transparency engenders contextual trust. Even if the data are not completely public, trust can be enhanced if the data are reviewed by groups other than those who generated and directly use the data. In some cases, however, making the data available can be difficult, especially if the data are voluminous — for example, driving data from autonomous vehicles. Complete openness also may inhibit physicians' ability to talk openly with each other about mistakes they have made.

The airline industry has created mechanisms for anonymously reporting near misses, which gives industry and regulators better data. Robotics could help flag near misses after the fact to gain more information about safety, though this would require regulatory changes to promote such reporting. This reporting culture is already well embedded in parts of medicine, so it is possible.

When the record of a robotic procedure is preserved, it has to preserve more than just what the device did. It needs to record the state of the entire system, including its human components. It also may be necessary or advisable to keep more data than FDA requires, so that data are available when issues or new questions arise later.

Institutions, companies, and countries could learn a great deal from each other. Sharing information could reduce duplication and accelerate innovation. Other countries are having the same conversations, and something can be learned from them. Countries also can act as independent laboratories, so that innovations tried elsewhere can be imported, and vice versa. Differences between countries and between younger and older people could broaden the conversation. In addition, training programs can address data issues by teaching people who generate, analyze, and store data how to do so appropriately.

Recommendations Related to Collecting and Using Data

Create *a stratified access system analogous to the military security system with different levels of protection. For example, some information could be exchanged that does not fall under HIPAA provisions, while information that is subject to HIPAA would require a higher level of authorization and protection.*

Establish *new legal and regulatory provisions to ensure that data are not misused. For example, insurance companies need HIPAA data to make payment decisions; should that information be sharable among companies? To what extent should patients be allowed to access medical or financial information?*

Standardize *data collection and curation where possible. For example, agreed-upon data fields and the collection of normalized data could enhance progress.*

Encourage *exchange of information among countries to reduce duplication and accelerate innovation.*


The use in the civilian sector of robots originally designed for military purposes could have major effects . . .
POLICE DO NOT CROSS

3. Focus Area: **Military Robots in the Civilian Sector**

on civil society. One of the Halcyon Dialogues on Robotics discussed this issue in depth as an example of the broader trend of roboticization in society. Participants discussed the broad issues of regulation, accountability, and trust, and like the group that discussed medical robotics, they developed recommendations in each of these issue areas.

Possible Uses of Military Robots in the Civilian Sector

Robotic devices have been used throughout the military. They are employed in border patrol, homeland security, and emergency response. They perform such activities as bomb disposal; precision targeting and strike; biological, chemical, and nuclear detection; transportation; reconnaissance; early warning; search and rescue; damage assessments; mapping and asset tracking; and humanitarian assistance. They offer versatility, persistent functionality, the capacity to reduce the risk to human life, and the potential to contribute across warfighting sectors.

Many of the robotic systems currently used by the military were developed in the civilian sector. For example, the Boeing Insitu ScanEagle, a small, long-endurance, low-altitude unmanned aerial vehicle, was first developed to collect weather data to help tuna fishermen locate and track schools of fish. The ScanEagle was first used by the military in 2004 and proved highly useful for autonomous surveillance in the battlefield before being replaced by the Boeing Insitu RQ-21 Integrator.

Recently, the military has created its own autonomous systems that may have applications in the civilian sector. For example, a partnership between the Kaman and Lockheed Martin corporations resulted in the K-MAX Unmanned Multi-Mission helicopter developed for hazardous military missions. It can be used to deliver supplies to the battlefield and, in civilian applications, can assist with chemical, biological, or radiological hazards. In June 2015, Kaman announced that it was restarting production of unmanned K-MAX helicopters after receiving ten commercial orders. Currently, inquiries are being made by representatives of firefighting, logging, and industry transport organizations.

Another example of a military technology that has transitioned to the public sphere is the Global Positioning System (GPS), which is used by robotic systems in many sectors. Similarly, semi-autonomous and fully autonomous automobiles probably would not yet exist without the impetus of the Defense Advanced Research Projects Agency (DARPA). Unmanned vehicles with quad rotors, humanoid robots, and enhanced exoskeletons, all of which have military applications, could find uses in a wide variety of civilian sectors.

LAW ENFORCEMENT APPLICATIONS

Initially, law enforcement will be a major user of transitioning technologies. Current uses of robots in law enforcement include explosive ordinance disposal and bomb squads. In 2016, Dallas police used an explosive device attached to a tele-operated vehicle for the first time to kill a suspect in a sniper attack that killed five officers.

Additional uses of robotic technologies in law enforcement include reconnaissance systems to supplement the closed circuit televisions that already cover most big cities. Autonomous or telepresence robots could patrol dangerous areas of the city while officers at headquarters monitor multiple systems at once. Virtual reality training could enhance officers' observational skills. Exoskeletons may allow officers to reach an environment more quickly and safely. Voice analysis, data mining, and machine learning technologies could enable rapid and accurate assessments of threat. Sensors on an officer or robot may be able to detect whether a gun is real or fake. Robot technology could be used for crowd control and for prison management.

Initially, law enforcement will be a major user of transitioning technologies.

OTHER APPLICATIONS

Robots designed for the military have many possible applications outside of law enforcement. Unmanned aircraft systems performed mapping after the 2015 earthquake in Nepal, unmanned surface vehicles have assisted in the water rescues of fleeing Middle Eastern refugees, and robotic systems have been deployed to assist in wildfires, flooding, and other natural disasters. To cite a specific example of a future potential use, new robotic devices and drones could autonomously survey wildfires, contain them, and mitigate human danger. In general, the potential for military technologies to transition to the civilian sector, and vice versa, is essentially unlimited.

However, the use of military robots in the civilian sector will need to be carefully considered. Military robots typically have been designed to perform specific tasks; in a civilian setting these tasks may have a different nature. In the military, situations are more black and white, one side against the other, even though situations can be extremely complex. The civil sector has far more shades of gray, with more opportunities for error and less concrete ways of determining success. In the military, decisions must be made that weigh the goal with acceptable costs and collateral damage. Police officers, for example, make decisions in a different manner due to the complex and often ambiguous circumstances faced in the field.

TRUST

As observed in the previous chapter, trust has many dimensions. It can include technical trust, operational trust, personal trust, and so on. Just as some people will never trust banks, some people will never trust robots. But other people's trust can be earned.

In many cases, the public will have to perceive that the performance of an autonomous robot is better than human performance in terms of reliability and safety to trust that technology. With automobiles, a commonly used metric is the statistic of one fatality for every 100 million miles driven by a human. Developers of autonomous vehicles will not be able to test that many miles without modeling and simulation. The University of Michigan, for example, built a track to test emergency situations and prove that the safety of an autonomous car can be improved through repeated testing under controlled circumstances. These results could be applied on a larger scale. Eventually, however, real-world test cases will also be necessary. A simulation is only as good as the model it is using, and people will need hard evidence to believe in a system's reliability.

Maintaining trust will require that margins of error in programming and algorithms shrink. In the early 21st century, an acceptable standard of error for a standard luxury car was 14 to 20 errors per 1,000 lines of code. The current standard is about 0.3 errors, and in 15 years this standard will probably be 0.001 errors. But in vehicles with hundreds of millions of lines of code, this standard could still result in hundreds of errors. When errors are found, software will need to be updated, which requires deciding how updates should occur, who should have the authority to change automated vehicle programming, and how testing should be conducted.

To maintain trust, laws and regulations must be enacted and enforced in a fair and reasonable manner. If laws are ambiguous, they will have to be reassessed, clarified, and commonly understood. If people feel that their civil liberties are being violated or if a robot causes harm to a person, public trust will evaporate. Just a single negative interaction among a sea of positive ones can change attitudes. The public will make threat assessments regarding such issues as hacking, cybersecurity, and safety and weigh threats against benefits to decide on the use of robotic systems.

Who delivers information and how much is received also help determine trust. Today, cellphones are endless sources of information. Even if this information is biased, it can vastly influence public opinion. Media messages also have a large effect and can cause public perceptions to differ by region.

Comfort levels with autonomous systems are different for different populations based on demographics. Adults are more familiar with robots now than they used to be because their children or grandchildren are learning robotics in school. Investments in education and training for all demographic groups may be as crucial as investments in technology in engendering trust. Perhaps different messengers or messages could be used to reach different populations. For instance, assurances from a NASCAR driver may be the most effective way to develop trust in autonomous vehicles in some parts of the country.

Receiving feedback from the public is as important as education. As one example, an estimated 260 million people play videogames in the United States. New technologies employed virtually with large cohorts of videogame players could generate enormous amounts of data. With this kind of testing, developers of a technology could assess what players understand, set up a chatroom for player feedback, and test policies through the games. To take another example, simulators in law enforcement are today used to expose civilians to the environments under which law enforcement officers operate. One such initiative, for instance, lets citizens engage in a simulated shooting incident, after which they provide feedback for both sides. Similar virtual environments could test how people drive or respond in emergency scenarios to develop better robotic technologies.

Trust in robotic technologies will initially and primarily be built through the teaming of people with autonomous systems. Trust of robots will grow gradually and through small steps, with small and repeated transactions leading to more significant ones. Such partnerships will enhance public understanding and acceptance of robotic and autonomous systems, their capabilities, and their proper uses. The primary purpose of a robot is to remove burdens from humans, and the message that robots are helping people and not hurting them will allow for their wider application. These interactions also will constitute part of the education process necessary for full integration of these systems into society.

These kinds of partnerships between humans and robots also will facilitate public trust in law enforcement. To the extent that people understand that police robotic systems will not be used by themselves, that humans will always have preeminence and control over these systems, and that these systems will only

use an appropriate level of force, they will be more likely to trust these working partnerships, and eventually the robotic technologies themselves.

In law enforcement, officers are valued because of their ability to make quick and accurate decisions in life-threatening situations. While robotic systems can and have been successfully used in these environments, use of these systems also could have disastrous results. When a robotic system is successful in its mission, it tends not to receive acclaim, but when that same system is used with harmful results, a public outcry can force it to be reexamined. As with police officers, robotic systems have the potential to make mistakes in "judgment." Yet just a single mistake can destroy communal trust.

As an example of how trust can be lost, a project in Baltimore conducted a pilot test of surveillance technology including aerial surveillance in solving mysterious crimes. However, the project failed to engage the community and to establish an appropriate privacy agreement, and a public outcry jeopardized its mission. A community must consent to be governed, just as it must consent to deploy robotic technologies in useful ways.

Robotic systems can collect vast amounts of information. How much of these data should be collected and how they should be used must be determined and agreed upon by all parties. Information also must be collected and analyzed on robotic systems, including their failures and successes. For example, the number of lives saved, not lost, through these technologies will be a salient statistic for the public, though this statistic is difficult to quantify.

The public may not appreciate when the intricacies of policy making. Rather, people want to know that policy makers are protecting their interests. Today, a portion of the population does not feel that way and does not trust the policy-making process. Trust will need to be built into the process at an institutional level.

> Trust of robots will grow gradually and through small steps, with small and repeated transactions leading to bigger ones.

The relationship between the public and law enforcement is also changing. Law enforcement has typically asked communities for their unconditional trust, but the expectations of communities are changing. Citizens now expect to recognize and understand the tools officers use to keep them safe, which will require more transparency at both the state and local levels of government.

Robots also can be programmed to communicate with humans in ways that engender trust. If a self-driving car can explain why a certain action was taken, the passenger's understanding might lead to greater trust. The opposite is also true: a lack of communication between an autonomous system and a passenger might lead to a lack of trust. In the military, users are well trained, so they might not require this level of communication. But civilians generally do not have this level of training and will need instruction to gain confidence in themselves and in robotic systems.

In emergency situations, the public must trust or rely on local emergency services. But in daily life, people will choose to use or not use robotic systems, and this choice depends on their sense of reliability and safety in these technologies. Value is created through a combination of need and trust.

Overall, trust in institutions has declined precipitously in recent years. The 2015 Edelman Trust Barometer analyzed the relationship between trust and a country's ability to innovate. The results showed that the more institutional trust a country has, the higher its ability to innovate. These results suggest that trust scaffolds and partnerships must be formulated to ensure the continued development of robotic systems.

Recommendations Related to Trust

Foster *interactions among the users, developers, and producers of robotic systems so that all parties understand the functions and limitations, imitations of these new technologies.*

Pursue *public education to foster widespread understanding of the capabilities and limitations of these systems. Continued dialogue will help enable informed and collective decisions.*

Introduce *new systems deliberately and with options for implementation that make the transition more manageable. An example would be giving autonomous vehicles a fully autonomous mode for freeways but a more interactive mode in highly populated areas.*

Address *fundamental ethical and regulatory questions earlier in the process of transitioning military robotic systems to the civilian sector.*

Accountability

Accountability has received considerable attention in the public sphere, especially in law enforcement, due to the highly publicized and protested string of recent police shootings. This focus could shift to robotic technologies. For these systems to be successful, firm and just accountability and liability practices must be in place.

Different liability issues apply to different professions. Both physicians and law enforcement officers take an oath, but they are treated differently. A patient enters into a consenting agreement with a physician. Comparable agreements do not exist with police interactions.

Liability considerations will influence how companies release technologies. For example, if insurance companies are liable, they could allow personalization of a device but set the policy's price accordingly. If a manufacturer is liable, customization is less likely so that all devices can be sold at a uniform price. Accountability will incentivize designers, manufacturers, and distributors to minimize errors. Already, many insurance companies offer safe driver discounts for autonomous and semi-autonomous vehicles if the user agrees to provide all the data from a car.

As the development of robotic technologies accelerates, one pressing question is whether a code of ethics can be created to encourage accountable and trust-building development of these systems. Ethical values could be embedded into the design, implementation, and use of these technologies, depending on the technology, its risks, and existing laws. At the level of design, designers and engineers could have a generalized toolkit that enables them to consider ethics. Ethical commitments also will need to be considered on a case-by-case basis and for particular uses in particular contexts. Such assessments may be more appropriate at the community level than the state or federal level.

Implementing ethics at the design level raises challenging questions. Who has the authority to ensure that these issues are considered? How will these issues make their way into public consciousness? The IEEE Global Initiative on Ethics in Autonomous Systems is currently working to develop guidelines for the design and implementation of such systems. An analysis of current and upcoming technologies in the military could help determine which systems are applicable to the public sphere. Legal experts also will need to identify ambiguities in the law with regard to robotic technologies.

Risk management strategies will be needed to manage the inherent uncertainties that arise from robotic technologies. For example, the Center for the Study of Existential Risk at the University of Cambridge is working with industry and machine learning experts to develop strategies for managing uncertainty in situations with the potential for dire consequences. Similar efforts will be needed to address other issues that continue to be surrounded by uncertainty.

Recommendations Related to Accountability

Provide *a solid and context-specific rationale for the introduction of military robotics into the civilian sector. Such a rationale will provide both a framework for accountability and a way to build trust.*

Use *broadly based analytics to evaluate the implications of potential applications of military technologies. Middle-sized organizations will probably be the first to adopt these new technologies. Perhaps a suburban law enforcement agency with fewer officers but a higher budget might be a good candidate for the initial introduction of a robotic infrastructure in law enforcement.*

Invest *in organizations and partnerships that can bridge the gap between innovation and commercialization. Today, organizations that can take a nearly ready technology, fix its bugs, make sure that it is highly reliable, and transfer it to someone who might use it are lacking. Teams of people could be developed who invented the technologies, with support from advocates who emphasize the benefits of the widespread application of these technologies.*

Regulation

Ultimately, the application of these technologies will depend on a variety of factors, including their regulation, cost, ease of use, and reliability. Regulation has both top-down and bottom-up components. The top-down approach looks at ways to modify existing laws to capture important values. The bottom-up approach seeks to establish legitimacy by getting people involved and having them contribute to procedural decisions at the local level. In either approach, the first step is to determine governance. Laws exist at the federal, state, and local levels. In addition, lawyers determine what is acceptable through lawsuits where laws and regulations have ambiguous applications.

Regulations and proposals to require the registration of drones suggest some of the issues that may be involved. Different states may regulate robotic technologies in different ways. For example, Virginia has historically been a leader

in airspace surveillance, while Alaska is interested in using robotic technologies to detect oil spills, track caribou, and discover breaks in the Alaska pipeline. Federal innovation can often result from the ability to launch experiments at the state level. Collaboration between all sectors of government will be crucial in this process.

People who are regulating robotic systems, whether at a federal, state, or local level, have to be trained on how the systems operate and the issues involved. Even more important, legislators will need to learn more about system capabilities and limitations to establish lasting, equitable, effective, and safe laws.

Regulations require consistency. For example, social norms exist in driving, like driving slightly over the speed limit or taking a rolling stop through an empty intersection. Should a self-driving car have to come to a full stop at every stop sign? If it gets rear ended when it comes to a full stop, how should the situation be handled? Who pays for tickets against self-driving cars? If a car is driven by a human, the state government generally regulates it, but if it is driven by software, the federal government regulates it, and laws are typically enforced by local officers. While guidelines exist for human conduct, guidelines for robotic conduct have yet to be created.

Customer involvement, feedback, and understanding will be crucial to a successful transition, but strategies to accomplish these goals must be determined. Possibilities include training or certification of users and manufacturers. Given that it is difficult to build a certification organization from scratch, perhaps an existing organization could be enlisted to provide such certifications and gain consumers' trust. The Underwriters Laboratories (UL) organization certifies many kinds of equipment, and the Consumer Reports organization judges equipment and its uses. Focus groups have shown that people prefer the idea of technology being assessed by independent evaluators rather than the developers.

People also prefer the development of protocols independent of a specific manufacturer. An organization like the National Institute of Standards and Technology in the Commerce Department that has been developing standards and testing products for more than a century could develop standards for robotics development. However, the development of standards in such areas as testing autonomy is difficult because results are never the same even in slightly different environments.

Another common model used in governing emerging technologies involves professional standards organizations, where industry experts come together to discuss how standards should be maintained and what regulations should be in place. This process typically does not involve the government. Rather, it involves developing a consensus in the private sector for the rules to be upheld. (Chapter 5 discusses regulatory issues in more detail.)

In law enforcement, prevention is the number one goal. The application of new technologies in any sector cannot threaten public safety. Restrictions and standards must be in place to protect human life. Metrics for effective policing can help determine the effectiveness of robotic technologies applied in the field.

Recommendations Related to Regulation

Create *a federal interagency coordinating body to further the development of legal, regulatory, and ethical decision making. Such a body could have both institutional buy-in and some level of independence in shaping how robotic technologies are governed.*

Look *to other communities, such as the medical community, for lessons in developing regulatory systems. An absence of legislation could force decisions to be made in the courts, which could result in arbitrary or uninformed rulings. Too much regulation could harm the innovation process.*

Make *use of national laboratories in transferring robotic technologies from the military to the public sector. These laboratories can take research done with a narrow purpose, further pursue it, and adapt it for the field.*

The Transition Process

Transitioning military robotic technologies into the civilian sector raises many important questions. What is the goal of the transition? What are appropriate ways for governments to use military-derived platforms? Who will be accountable for system malfunctions and injuries to human life? What are appropriate safeguards to protect the public from misuse? Should anything be declared off limits?

The Department of Homeland Security has invested in the transition of military robotic technologies to civilian use. However, the Department of Justice has devoted much less funding and attention to the issue. Greater emphasis and investment in the transition process will be necessary to make progress on outstanding issues.

When GPS technology moved from the military to the civilian domain, it was not a wholesale transfer. The position data received from a civilian GPS receiver are different compared with the data from a military GPS receiver. Such examples could serve as both positive and negative precedents in determining how to deal with the integration of these new technologies into society.

Given that determining "peaceful" uses of robotic technologies is highly subjective, a better goal is to minimize risk. Risk is the likelihood of a negative outcome and the severity of that outcome. Thus, risk can be lessened by reducing the likelihood of something happening and by introducing factors to reduce the severity of that thing happening.

Once robotic technologies are ready to be introduced, a testing environment is needed that is suitable for trial. Optimized outcomes for these systems should focus on end users, and the first demonstrations of these systems must be geared toward their specific applications. Addressing the needs of users can help identify the situations in which the use of these technologies will provide maximal benefits.

Regarding privacy, important questions are whether these technologies can collect information on their users and whether they can be hacked. The emerging use of big data in new technologies also has the potential to create risks of bias or discrimination, either implicitly or explicitly. (The uses of big data in a medical context are discussed more extensively in Chapter 2.)

Resistance to unmanned systems in the military is both cultural and generational. In the military, cultural shifts tend to coincide with generational shifts. The younger generation will likely be more accepting of new technologies because their trust in those technologies will be built over time.

The military does not currently use lethal autonomous weapon systems (LAWS). People in the military still widely agree that a human should have to make the decision to kill another human. However, should LAWS be employed in law enforcement? Law enforcement seeks to promote law and order and protect lives. The translation of robotic systems into law enforcement needs to be guided by these goals.

Cultural shifts tend to coincide with generational shifts. The younger generation will likely be more accepting of new technologies.

In the third meeting of the Halcyon Dialogues on Robotics, participants discussed several broad issues . . .

4. The Implications of Robotics for Work and Social Justice

beyond the specific contexts examined at the first and second meetings. They looked at automation and the workforce, at the level of individuals, and entire industries and nations. They also addressed how design and implementation can serve the greater good, including the ways in which priorities for technical development can be optimized to ensure that these technologies yield maximum benefits. And they looked at education and training as ways to both reduce the disruption caused by robots and cope with the disruption that inevitably will occur.

Robotics and the Workforce

As part of a broader trend toward the automation of jobs previously performed by humans, robots could reshape the world of work. The effects to date have been most visible in manufacturing, but robots eventually could do many of the jobs currently done by people in many sectors of the economy, including service sectors.

Greater use of robots will both eliminate and create jobs, though exactly how these negative and positive effects will be distributed remains highly uncertain. In many cases today, robots cannot replace highly skilled workers, and this will continue into the future. However, robots may be able to augment the skills of workers, serving essentially as "co-robots." Such robots could enhance the skills of less skilled workers, whether by direct intervention in jobs or by training, and they could make more skilled workers even more productive. Such robots could be especially valuable in situations where skilled workers are hard to find or are aging out of the workforce.

One argument in favor of robots replacing humans is that they can handle jobs characterized by the three D's — dull, dirty, and dangerous. For example, some underground mining jobs may be too dangerous for a human but could be done by a robot; some agricultural jobs could provide opportunities for robotic replacements; and robots could deal with hazardous materials released as

sea level rises along industrialized shorelines. However, even dull, dirty, and dangerous jobs are still jobs. Workers may have been doing them for a long time and command higher wages because of their skills, or those may be the only jobs available.

Many companies will need to automate to remain competitive. If they do not do so, the jobs these companies provide and the jobs associated with those companies will be lost. If greater use of robots increases productivity, more resources will be available to spur spending and employment growth in general. Robots also will expand many job categories and create entirely new kinds of jobs — for example, new jobs that are integrated with tasks performed by robots.

Through these and other mechanisms, robotics will change the nature and numbers of jobs throughout the economy. To take a negative example, the expansion of robotics in one sector may cause the loss of ancillary jobs. If some portion of the more than 3 million truck drivers in the United States were to lose their jobs because of greater use of autonomous vehicles, the people who serve truck drivers at gas stations, rest stops, and motels could lose their jobs

> Robots eventually could do many of the jobs currently done by people in many sectors of the economy, including service sectors.

as well. However, the pace with which robotics advances will help determine the social effects of robots. Rapid replacement of human jobs could create greater dislocations than a slower replacement. The jobs of many truck drivers may eventually be lost as vehicles are automated, but whether that process occurs over 5 years or 20 will make a big difference.

The application of robotics will also create many jobs. For example, the use of robots in social and personal environments is likely to have a large economic multiplier effect because so much of those environments will need to be modified to accommodate robots. Homes, vehicles, offices, stores, and many other settings will need to be retrofitted and outfitted with new devices. This is true not only in high-income but in low- and middle-income countries, which also will evolve in coming years to accommodate much greater use of robots, including uses that do not exist in high-income countries.

The effects of automation will vary greatly by sector, by region, and by country. For example, robots could be competitive with low-skilled jobs overseas jobs. If clothes can be made by robots, perhaps with increasing levels of personalization, they no longer will need to be made in low-income countries, giving high-income countries further advantages. On the other hand, countries with less experience with robotics could leapfrog other countries, especially if their educational systems build expertise and enthusiasm for robotics.

A major question is whether robotics will increase or decrease inequities among groups of people or nations. Productivity growth could benefit all workers, but in the recent past most of these benefits have been going to higher income workers. The returns on investments in robotics could flow to a relatively narrow portion of the population rather than being broadly distributed. Whether growing inequality could heighten social instability remains to be seen.

Countries will face many difficult public policy questions in responding to expanding uses of robots. Could robots be used in ways that would employ more people, especially people with lower skill levels? In general, automation will be more disruptive in places that have relatively anemic public policies to respond to job displacement.

Beyond the workplace, the expansion of robotics could alter social structures, such as gender roles, as they change what people do in workplaces and at home. Science fiction writers can help explore possible futures, as they have done in the past.

Recommendations Related to the Workforce

Support *research to examine the effects of robotics on the workforce and on the broader society. Fundamental public policy questions include: What is the pace of change? Who is being affected by change? How are they being affected? Where is change happening? As one example of this kind of research, the National Science Foundation has been exploring the possibility of funding research on the human-technology frontier, including robotics, as part of a new program on convergence in science and technology. Research funded under the program would investigate the social, behavioral, and economic impacts of new technologies, yielding a holistic view of their impacts on society as a whole.*

Design and Implementation

The design and implementation of robots are critical factors in how they will be used and whether they will be accepted. In some settings, for example, just the word robot can generate trepidation. But if robots are designed and presented as tools, people will be less likely to fear them or see them as threatening. Older adults, for example, are not necessarily afraid of technology. If it is valuable to them and they can use it, they will do so. Otherwise, they will not waste their time.

Reliability and predictability are critical in robots that work in the social and personal environment. If a television works only 80 or 90 percent of the time when people turn it on, they will grow frustrated with it. That is one reason why some people have trouble with smartphones — the same inputs do not always generate the same results.

The social expressivity of robots can be an important factor in their acceptance and use. Robots that are aware of the social, cultural, and gender cues of users can be much more responsive. Similarly, a robot that expresses disappointment or remorse when making a mistake is likely to receive a more sympathetic reaction from a human. On the negative side, if robots are perceived as being social entities, people could divulge private information to robots, or they could be deceived by robots.

Trust is an important element in design and implementation, and trust, as discussed in the previous chapter, has many dimensions. Can someone driving an autonomous vehicle trust that vehicle to make good decisions when facing difficult choices, such as choosing between harming a pedestrian or hitting another vehicle? As robots learn from their interactions with humans and other machines, will humans be able to continue to trust the decisions they will make? Can robots be trusted to keep personal information private, especially when that information is stored in the cloud so that comparisons can be made and patterns detected?

Preserving human dignity is another factor in design and implementation. Can the human dignity that comes from work, health, and positive relationships be preserved in a world of machines?

Overpromising and underdelivering can harm a nascent industry. If people's expectations are not fulfilled, a backlash can develop against the industry from both the public and policy makers. People can become disillusioned if robotic health care workers do not materialize. They can become disillusioned if robots make mistakes. People in the robotics industry have a responsibility to set and manage expectations.

At one meeting, participants discussed the possibility of establishing ethical guidelines for robots that could be embodied in design and implementation decisions. For example:

- Robots should not be autonomous killing machines.
- Robots should abide by laws.
- Robots should be good consumer products and should not be described in deceptive terms.
- Robots should not be able to manipulate vulnerable users.
- Robots should be as least expressive and manipulative as possible.

Such guidelines inevitably raise difficult issues. For example, some argue that it is impossible to ban autonomous weapon systems, adding that many semi-autonomous weapon systems exist. Just as sophisticated electronic processing was put in bombs, it will be put in robots, including robotic weapons. If some countries forego autonomous weapons, other countries will gain an advantage by adopting them. A counterargument is that countries have many shared interests and do not need to engage in intense competition on every front. Governments will regulate robotic technologies (as discussed in the next chapter), just as they regulate airplane and automobiles, and governments can agree to outlaw robotic weapons, just as they have outlawed chemical and biological weapons.

> Can the human dignity that comes from work, health, and positive relationships be preserved in a world of machines?

The users of robots are the experts on that use, not the designers of robots. Chapter 2 pointed out that robots could serve many valuable purposes with older adults, such as helping them move around, cook, get dressed, and stay in their own homes. But older adults who use robots also will be the experts on how to improve those devices and their functions. By engaging with users, designers can take advantage of their expertise. In effect, the users become members of the design team.

Involving the users in design helps create a bottom-up perspective. It reveals what a person or community needs, which provides a way to build value and sustainability into a robotic system in a particular context. Participatory design that involves people with a diverse array of skills and backgrounds can solve problems in ways different to those that would have been the case otherwise. It can reflect local needs, capabilities, and values, so that robotic technologies serve the people who they are designed to help. Technologies that reflect local needs are also less likely to break or not be useful, which is a problem with many current technologies and has the effect of souring people on technological innovation.

Recommendations Related to Design and Implementation

Foster *continued dialogues among people representing many points of view to create more representative design processes.*

Develop *and implement best practices to establish and maintain trust. An example involves the transparency of data flows. If a camera is in a house and is generating data, that process needs to be visible. Another best practice is to have multiple trust vectors. Thus, if a machine is approaching someone, that person should have at least two indications of what can be expected, whether conveyed through color, lights, shape, movements, speed, proximity, or other indicators. In this way, robots can leverage the expectations embedded in human biology to serve human purposes.*

Development *and implementation of a familiar design library can be leveraged to generate trust, leading eventually to a vocabulary of design that maintains trust. This aspect of design is insufficiently articulated in current robotic design and needs to be included from the outset, not added on later.*

Use *a third party authentication group to enhance trust. For example, a global robotics ethics and efficacy authentication process could generate trust, whether conducted through governments, social media, or industry groups. By being international, such an authentication process would include diverse cultural perspectives on such issues as stereotype perpetuation and empathy.*

Education and Training

The age of robotics will accelerate and intensify many of the trends in education and training initiated by the information age.

Today, many employers are having trouble finding workers with the technical skills needed in robotics. Such workers often are trained in community colleges, technical schools, certificate programs, or apprenticeships. These programs also can help retrain workers whose jobs have been lost because of economic or technological changes.

A more responsive educational system can ameliorate the negative consequences of automation. In the future, the majority of jobs could involve some aspects of engineering and computer science, and education will need to move in that direction. Blue collar and white collar jobs and other traditional formulations will lose their meaning as roles and responsibilities are redefined and blended.

As the economy changes, people will need scientific and technical skills to fill many jobs, but they also will need what have been called soft skills, such as the ability to learn quickly, to communicate effectively, and to work in teams. One way to create such combinations of skills is by integrating the arts and humanities into science, technology, engineering, and mathematics (STEM) disciplines, creating an interdisciplinary curriculum sometimes referred to as STEAM education. With robotics, for instance, combining engineering with the arts and social sciences could lead to designs for service robots that address not just physical needs but emotional needs. Such robots can use facial expressions, gestures, and physical movements to communicate, making communications more complete and efficient. The result can be robots with which people feel more comfortable and receptive.

This kind of hands-on training, which used to be developed by shop and home economics classes, is reemerging in schools through such programs as maker spaces and Project Lead the Way, which provides K-12 students with applied learning experiences to enable them to thrive in college and in careers. As robotic systems are integrated into everyday life, the concept of collaboration with robots will evolve. Teaching these skills at an early age will equip generations with the tools needed to handle the new dynamic between technology and human life.

Another approach to building combinations of hard and soft skills has been through afterschool programs such as the FIRST (For Inspiration and Recognition of Science and Technology) robotics programs, which have involved hundreds of thousands of students in designing, building, and operating robots. Such activities could be more thoroughly integrated into the K-12 curriculum, as could such subjects as engineering and computer science.

Pedagogies that are culturally responsive can increase the interest and engagement of students. When students in underserved communities, for example, are introduced to issues that are important to them and to their communities, they can see and understand how technology can be used to address these issues. The same argument applies to students with disabilities, who have expertise that can benefit much broader groups.

Online training is becoming both more versatile and more common. It can be done anytime and anywhere, encompasses many more topics than in the past, and can be directed toward people with a wide variety of backgrounds. Virtual reality, telepresence, and other technologies could also help with retraining.

Robots eventually could do many of the jobs currently done by people in many sectors of the economy, including service sectors.

An option being pursued by governments in some other countries is to subsidize the expenses an individual incurs each year on retraining, which emphasizes the message that people are responsible for their own retraining and need to think continually about how to update their skills.

Nevertheless, retraining is not an option for everyone, and in some places the term itself is disparaged because of unsatisfactory results in the past. Some industries and communities are going to continue to lose jobs, and not everyone in those industries or communities can learn the skills to acquire a job that pays as well as a job that has been lost. Better public policies will be needed to deal with the inevitable human consequences of roboticization.

People will need training to work with robots, but robots also will become easier to use. In this regard, robots are several decades behind computers. When anyone can use a robot in the same way that they use a cell phone, robots will have even greater potential impacts.

People can have a visceral fear of robots, grounded more in science fiction than current realities. The status and implications of these technologies must be communicated to the public in some sort of intuitive, instinctive manner. Developers will be responsible for determining the facts to convey to the public, but a different approach is needed to connect the public to this information. Technologists, academics, and researchers could provide facts and more solid timeframes to the public to lessen the hyperbolic way in which these systems are often discussed, both in the media and by the public.

Machine learning and data analytics can help companies capture and take advantage of the human capital that they have. They can identify workers who can more easily be trained to acquire new skills, even those who have worked with their hands their whole lives. Such workers may have characteristics like adaptability or communications skills that data analytics could reveal. This can bridge the gap between what companies need and what their existing employees can do.

The scale of educational reform needed for the robotics age may be on the scale of the GI Bill after World War II, though more nuanced policy tools are available now than was the case then. The formation of an organization like DARPA in the Department of Education could help bring about the innovative tools and methods that will be needed to adapt to the advance of robotics.

Recommendations Related to Education

Apply *the methods and tools associated with robotics, such as machine learning, augmented reality, and virtual reality, to radically improve the quality of education.*

Institute *modern versions of the industrial arts in the K-12 curriculum as robotics leads to more mechanized workplaces.*

Create *an Advanced Research Projects Agency - Education (ARPA-ED) to explore innovative and technology-intensive approaches to education at the K-12, undergraduate, and graduate levels.*

The final Halcyon Dialogue on Robotics turned to matters of law, litigation, and liability . . .

5. The Intersection of Robotics and Policy

at the intersection of robotics and public policy. After a discussion of issues associated with minimizing the risks and maximizing the adoption of robotic technologies, the group considered how best to govern and oversee these devices in the dawning age of robotics.

Minimizing Risk and Maximizing Adoption

Minimizing the risks of emerging robotic systems will be a critical factor in maximizing their adoption. The public is less likely to protest low-risk technologies than high-risk technologies. They also are more likely to support human-machine teams than a full replacement of humans with robotic systems.

Building trust in robotic technologies requires testing and evaluating these technologies. Currently, hardware is tested in a comprehensive and deterministic manner. All the different modes are evaluated to make sure that the system operates as expected. Devices are typically operated to failure to understand their performance.

However, exhaustively testing the whole realm of possibilities in which an autonomous system can operate is impossible. Developers know situations in which the device will definitely work, and they know situations in which it definitely will not work. Areas in which its operations are uncertain must be determined and fully tested. Determining these borders where robotic systems have the potential to fail will require trial and error, accidents, and incidents that will test the public's trust in these systems. Current liability standards will probably apply to the evolution of these technologies, and the accountability derived from these standards will reduce the margins for error.

As an example, cruise control is a commonly used robotic device in the civilian sector. Despite its capabilities, the driver is expected to pay attention and remains liable for damage. Eventually, as these expectations change, personal liability issues may transition to a governing body that tests the safety of these systems, or manufacturers may be held to a safety code that clarifies these types of liabilities.

As artificial intelligence becomes more sophisticated, testing and evaluation will have to involve not just lines of code but neural networks with millions of parameters. Resulting systems may be more difficult to characterize and assess for threat. Assumptions about the statistics, sensor data, environmental conditions, relational feedback, and user characteristics will all need to be incorporated into testing conditions to define boundaries for error.

Risk is also related to user interfaces. For example, someone driving an autonomous car may suddenly need to take control back from the car. But a driver needs time to reorient to the situation and to the potential for an accident. One question is how to have a human pay attention to what a computer is doing and then decide when to take over control. And does a human then assume responsibility from the computer when deciding to take control, or is the robotic system still responsible for any outcomes?

Engendering trust in robotic systems will be important to their integration into society, but overly trusting these systems also will create problems. Trusting robotic systems to perform beyond their capabilities or attributing personalities and relationships to these systems may encourage improper use. People have already made this mistake by overly trusting autopilots in semi-autonomous vehicles despite access to information on the vehicles' limitations. Some people also try to subvert and confound technological systems. For example, hackers may try to gain additional features from a system or break into other users' systems.

Modern autonomous vehicles heavily rely upon visual sensors, but soon they will evolve to include other sensors that are based on other types of feedback, which will enable redundancy in a system. For instance, a robotic device may separately receive and then integrate spatial and navigational cues through LIDAR, GPS, and an onboard, locally stored mapping system. For this reason, the reliability of robotic sensors will be as important as the number of sensors in robotic designs. However, manufacturers also must weigh sensory feedback with the cost and size of a product. For example, drones can be manufactured that are much safer because they have more redundancy and backup systems. However, the added weight of these additions exceeds Federal Aviation Administration limitations for safe use in populated areas.

Data mining has the potential to improve hazardous infrastructure. For instance, Google has catalogued most of the world's roads, and these data can be incorporated into autonomous vehicle feedback to improve safety. In addition, data collection, transmission, and storage by robotic devices will help minimize risk. Edge computing or preprocessing of the data collected by robotic systems in various contexts will be necessary to determine what data are worth saving and analyzing. Cars produce astronomical amounts of data that are impossible to capture even with an increase in storage and bandwidth. A large subset of computers will have to be dedicated to categorizing and eliminating data. Regulations may require systems to share certain data sets or delete others. The public fears a loss of privacy and security in data collection. These concerns must be weighed against the opportunities it provides.

A "one size fits all" approach will not work in creating a framework for data collection and sharing. Not everyone needs or should have the same access to the same information. Cars may receive one type of information to stay on the road, insurance companies may receive another about the driver's behavior, and a divorce proceeding might receive a third on a driver's destination. Systems must be sufficiently specific, varied, limited, and detailed to address these specific needs.

In addition, assumptions and practices vary from country to country. The European Union has aimed to create a consent-oriented system where users have the right to provide or withdraw their information while recognizing that different countries even within Europe have different sensibilities regarding data collection and use.

As robotic systems become more autonomous, humans will need to be able to understand the justification behind decisions made by those systems. Social behavior should be factored into this analysis, especially in considering human-machine teams. Foreseeable tampering of a product is only part of the complex array of factors that play into a product's liability analysis. Manufacturers are responsible for releasing products that do not create consumer or policy backlash.

Transparency in testing will encourage public trust and the adoption of new systems. A feedback loop among affected communities, developers, and other stakeholders will provide a more effective and supported integration of these systems into society.

Recommendations Related to Minimizing Risks and Maximizing Adoption

Require *of transparency for testing regimes so that the public is informed about testing practices.*

Incorporate *a diversity of voices in testing regimes, so that people of all backgrounds, ages, and experiences are heard and factored into the decision-making process.*

Design *robots so that communication to implement their behaviors comes through multi-sensory signals such as sounds, lights, and motions, similar to the way that car brake lights are used to indicate a reduction in speed. In this way, robots could indicate their current state to a user. They also could indicate their intentions, similar to the way that a car's blinking turn signal allows other drivers to anticipate its behavior. Modes of communication and definitions of robotic states will have to be invented.*

Governance

What is missing from current policies that will keep society from confronting the future of robotic technologies? As in many other areas, the regulatory questions to address depend on how robots are used and on the context of their use. In some situations, for example, regulatory certification will be necessary; in others, no formal certification process will be required. Some devices, like medical or automotive systems, involve both regulatory compliance issues and product liability issues. The military and government deal with fourth amendment and privacy issues, while the private sector deals with tort issues. Within the legal and regulatory system, physical harm, privacy harm, and property harm mean different things. All these different contexts will need to be dealt with separately on a regulatory level.

Governments have a range of regulatory tools at their disposal. At one extreme, they could ignore robotics and do nothing. At the other extreme, governments could ban robotics. Neither of these extremes is advisable. Achieving public trust and legislative action without stifling innovation and competition is a difficult task. New regulations may be needed; some existing regulations may need to be removed. Policy makers also have to determine the optimal timing for policies. Regulations created today can be problematic if they are outdated or do not apply in ten years. Sometimes, frameworks like voluntary best practices provide enough structure that additional policy is unnecessary. Allowing situations to evolve over time before creating legislation and regulations allows stakeholders to grasp potential directions and make decisions that are more sound.

Governments could encourage the producers of robotic systems to self-regulate, with government also playing a role. For example, the National Telecommunications and Information Administration has participated in the development of privacy best practices for unmanned aircraft systems. Governments could impose a penalty for breaking self-regulatory frameworks, as the Federal Trade Commission does with companies that do not conform to their privacy policies. Instead of prescribing a particular requirement for robotic systems, policy makers could encourage a variety of approaches and then develop a performance evaluation framework to determine whether a device meets safety standards.

The United States relies heavily on private standards development organizations, hundreds of which form and enforce industry standards. These organizations are in turn supported by larger systems to enforce standards and determine negligence. In some instances, private standards are so well articulated that they are adopted, at least in part, by regulatory agencies. The Occupational Safety and Health Administration, for example, has reproduced private practice policies in its own regulations. Often the government lacks the technical ability to make sufficiently informed regulations, so industries must be responsible for setting the standards themselves.

One complication is that the issues are inherently international. Robots can be operated remotely from anywhere in the world. Already, multinational drug

companies have research laboratories in different countries to conduct trials and tests specific to the local jurisdiction. How can national laws and regulations be respected in an inherently international setting?

Public policies also can encourage the growth and development of new technologies. The technologies that government agencies buy, fund, and use have the potential to influence public opinion and the market. For instance, the U.S. government's emphasis on electric vehicles caused a corresponding growth in the public sphere. In the past, policy makers have sometimes decided to limit liability in certain areas such as aviation, drug development, and nuclear power — for example, where market opportunities or requirements outweigh concerns about a technology. Governments can provide tax incentives or limit tax exposure. They also can support research, both on the technologies themselves and on the rules governing those technologies. The testing and evaluation of autonomous systems can inform funding priorities at the National Science Foundation and other research agencies.

> The creation of 'policies that transcend uncertainty' is a great challenge.

The many different contexts in which robots can operate make the development of universal standards difficult. Performance-based standards can be preferable to mandated standards that quickly become outdated. Governments can guide the development of a technology without creating a rigorous framework for its future. For example, governments can support studies that analyze possible pathways into the future and how to influence future developments. Increasing science diplomacy between countries can reduce uncertainty on an international level.

The creation of "policies that transcend uncertainty" is a great challenge. Uncertainty is unavoidable when considering future developments in artificial intelligence, autonomous systems, nanobots, and microsensors. Fostering innovation means doing things in ways that have not been foreseen. However, the regulatory framework need not be uncertain and could help robotics prosper as the field grows and develops.

Recommendations Related to Governance

Use *a variety of policy tools to support and encourage the responsible development and application of robotic systems, such as tax incentives, support of research and development, and purchasing subsidies.*

In *situations where producers of robotic systems self-regulate, determine the implications of failing to adhere to those regulations.*

Encourage *multi-stakeholder dialogues along with conversations arranged for particular purposes, such as establishing industry standards or best practices. Already, many government bodies facilitate industry standards, with companies that adhere to standards receiving liability benefits.*

Advisory Structures

For the most part, the United States' legal system will be able to handle the introduction of robotic technologies into society. The United States has a well-established body of laws and regulations that apply to robotic systems. If applied in a thorough but flexible way, these laws and regulations could help foster public trust of robotic technologies.

However, a much larger breadth of engagement is needed to encompass the range of people who will be affected by robotics. Without feedback from all types of people, manufacturers will make products that will not work for a substantial portion of the population. Increased engagement will allow for a system of accountability, input, and improvement.

Nearly every sector of society should have a voice in considering the oversight of robotic technologies, since all sectors and individuals will be affected. Specific sectors impacted by robotics that should take part in this discussion include the legal sector, labor, education, health, infrastructure development, environmental protection, national and international security, diplomacy, privacy, transportation, international trade, policing, agriculture, energy, diversity and inclusion, social justice, space, maritime, commerce, arts, and media. Involving all sectors will increase public trust. If members of the public perceive that only one sector of society has input into regulations, they will distrust the results. An increasing breadth of engagement from all parts of society can help put robotic systems to optimal use.Within government, some multidisciplinary and multiagency bodies already exist that can advise government regarding decisions about priorities, policies, and regulations. For example, governance of genetically modified organisms has taken shape within a broad multiagency governmental framework that has so far proven sufficient for handling the complexities of the issue. Robotics may require a similar interagency framework that can guide governmental decision making. The National Science and Technology Council within the Executive Office of the President could catalyze such a discussion.

More broadly distributed and deeper expertise can pay long-term dividends. Every year, AAAS places approximately 250 PhD-level scientists and engineers in government agencies for one to two years through its Science and Technology Policy Fellowships Program. Fellows bring their technical expertise to the agencies and return to their professions with increased political savvy. There are now 3,000 former fellows, many of whom remain in government service. AAAS also has promoted the idea of science diplomacy among nations and has advocated for creating science advisors in governments around the world. Specifically in the area of robotics, AAAS has started the journal Science Robotics to disseminate information about the field and to publish landmark papers in the field.

Competitiveness deserves to be a major consideration in public discussions of robotics. A high-level CEO from the robotics industry commented at the final meeting of the Halcyon Dialogues on Robotics that automation is happening no matter what. For policy makers, a major question will be whether the robot being used was produced in China or the United States. For the United States to maximize its influence and the benefits to its citizens, U.S. companies should provide the next wave of robotics. Various organizations publish reports on the competitive relations between nations, including industrial automation. But very little public discussion involves the United States' global position in robotics. Academia and government, as well as private companies, can help the United States maintain its competitive edge.

Recommendations Related to Advisory Structure

Create *an independent organization or augment an existing organization to provide assistance to the robotics sector in managing the risks and encouraging best practices. Ideally, this would be a global organization affiliated with existing for-profit or non-profit organizations. Organizations might also be domain specific. For example, A3, the Association for Advancing Automation, provides the industrial equivalent of robotics standards. There may be separate associations for transportation, logistics, and so on.*

Increase *fellowships in government to connect scientific and technical experts with policy makers and vice versa. The military is also working to develop cross-connections by expanding corporate fellowships that place military officers in companies for a year. Dialogues, internships, and externships among companies, nonprofit organizations, and government will allow better responses to the inevitable changes caused by robotic systems.*

Convene *a group of high-level leaders to survey the field of robotics with both breadth and depth. High-level leaders can convey information about the robotics sector more effectively than can developers. The Defense Science Board, for example, is a committee similar to the one proposed here, composed of senior-level people who cover the breadth and depth of the issue and communicate through campaigns grounded in productive discussions and visualizations.*

Additional **Reading**

The following reports provide additional information about the future of robotics and possible policy responses. This is not an exhaustive list, and many other sources of information are available.

Ad-hoc Industrial Advisory Group. 2010. *Factories of the Future PPP: Strategic Multi-Annual Roadmap*. Brussels: European Commission. Available at https://ec.europa.eu/research/industrial_technologies/pdf/ppp-factories-of-the-future-strategic-multiannual-roadmap-info-day_en.pdf

Computing Community Consortium. 2008. *A Research Roadmap for Medical and Healthcare Robotics*. Available at http://bdml.stanford.edu/twiki/pub/Haptics/HapticsLiterature/CCC-medical-healthcare-v7.pdf

Computing Community Consortium. 2009. *Roadmap of U.S. Robotics: From Internet to Robotics*. Available at http://www.us-robotics.us/reports/CCC%20Report.pdf

euRobotics aisbl. 2013. *Strategic Research Agenda for Robotics in Europe 2014–2020 (SRA)*. Available at https://ec.europa.eu/research/industrial_technologies/pdf/robotics-ppp-roadmap_en.pdf

NASA. 2015. *NASA Technology Roadmaps — TA4: Robotics and Autonomous Systems*. Available at https://www.nasa.gov/sites/default/files/atoms/files/2015_nasa_technology_roadmaps_ta_4_robotics_and_autonomous_systems_final.pdf

SPARC: The Partnership for Robotics in Europe. 2015. *Multi-Annual Roadmap for Robotics in Europe—Horizon 2020*. Available at https://eu-robotics.net/sparc/upload/about/files/H2020-Robotics-Multi-Annual-Roadmap-ICT-2016.pdf

Appendix **Halcyon Dialogue Session Participants**

Larry Albin, Deloitte Consulting
Colin Allen, Indiana University – Halcyon Dialogue Programming Committee
Ronald Arkin, Georgia Institute of Technology – Halcyon Dialogue Programming Committee
Robert D. Atkinson, Information Technology & Innovation Foundation (ITIF)
John Basl, Northeastern University
Jim Burch, Police Foundation
Jeff Burnstein, Association for Advancing Automation
Andrea Calloway, Towson University
James L. Cambias, Writer/Game Designer
Joanne Carney, American Association for the Advancement of Science (AAAS)
Howie Choset, Carnegie Mellon
Kevin Cleary, Children's National Health System
Brenda Cooper, Novelist
Rebecca Crootof, Yale Law School
Eric Daimler, White House OSTP (Fmr.) – Halcyon Dialogue Programming Committee
James Drake, The Hospital for Sick Children, Toronto
Peter Drivas, Honeywell
Pierre E. Dupont, Boston Children's Hospital, Harvard Medical School
Lisa Ellman, Hogan Lovells
Ed Finn, Arizona
Doug Gage, Technology Consultant
Amy Glasmeier, MIT
Nobuhiko Hata, Brigham and Women's Hospital
Bruce Heiman, K&L Gates
Rush Holt, American Association for the Advancement of Science (AAAS)
Masato Iehara, Mitsubishi Heavy Industries America, Inc.
Peter Kazanzides, Johns Hopkins University
Andra Keay, Silicon Valley Robotics
Ayorkor Korsah, Ashesi University College (Ghana)
Dorian M. Korz, FDA, Division of Surgical Devices
Margaret Kosal, Georgia Tech
Alex Kozak, X, the moonshot factory – Halcyon Dialogue Programming Committee
John W. Krakauer, Johns Hopkins University, School of Medicine
Sachiko Kuno, Halcyon
Junichiro Kuroda, Embassy of Japan
Cara LaPointe, U.S. Navy
Cori Lathan, AnthroTronix

Ashley J. Llorens, JHU/APL
Terah Lyons, OSTP (Fmr.)
Raj Madhavan, Humanitarian Robotics Technologies
Mitch Mandell, Thompson & Knight LLP
Sylvain Martel, Polytechnique Montreal (Canada)
Elena Messina, NIST – Halcyon Dialogue Programming Committee
Russell Moy, Southeastern Universities Research Association (SURA)
Robert O'Connor, National Science Foundation
Steve Olson, Freelance Writer – Report Author
Yasuo Osone, Hitachi, Ltd.
Taskin Padir, Northeastern University
Rajni Patel, Western University (Canada)
Joseph Persichini, Jr., Washington DC Police Foundation
Bill Powers, Potomac Institute for Policy Studies
David Rejeski, The Wilson Center
Elana Reman, K&L Gates
Deborah Runkle, AAAS
Veronica J. Santos, UCLA Mechanical and Aerospace Engineering
Bud Sparhawk, SIGMA
Haijun Su, Ohio State University
Leila Takayama, UC Santa Cruz
Russell H. Taylor, Johns Hopkins University – Halcyon Dialogue rogramming Committee
Dawn Tilbury, University of Michigan
Masahito Togami, Hitachi America, Ltd.
Rei Tsuchiya, Hitachi, Ltd.
Edward W. Tunstel, Johns Hopkins Applied Physics Lab
Shannon Vallor, Santa Clara University
Howard Wactlar, NSF — Information & Intelligent Systems (IIS)
Trista Wagoner, Science Robotics
Thomas Walker, Mitsubishi Heavy Industries, Inc.
Kai Wang, Johns Hopkins Applied Physics Lab
Darrell West, The Brookings Institution
Andrew B. Williams, Marquette University
George F. Wittenberg, Dept. of Veterans Affairs; University of Maryland
Ruth Wylie, Arizona State University – Halcyon Dialogue Programming Committee
Guang-Zhong Yang, Imperial College London
Dwight Yen, FDA, Division of Surgical Devices

Kate Goodall, Halcyon
Josh Mandell, Halcyon
Danielle Reed, Halcyon

Halcyon

Halcyon seeks and celebrates creativity in all forms and galvanizes creative individuals aspiring to promote social good. We bring together diverse groups of changemakers in art and social enterprise and provide a safe haven for their bold ideas to take flight. Halcyon offers an ecosystem of advocacy that encourages socially engaged creatives to learn, freely experiment, sometimes fail, and advance their talents and visions. In doing so, we foster new pathways to knowledge and resources, and help innovators transform their inspiration into impact.

AAAS

The American Association for the Advancement of Science is an international non-profit organization dedicated to advancing science, engineering, and innovation throughout the world for the benefit of all people. To fulfill this mission, the AAAS Board has set the following broad goals:

- Enhance communication among scientists, engineers, and the public;
- Promote and defend the integrity of science and its use;
- Strengthen support for the science and technology enterprise;
- Provide a voice for science on societal issues;
- Promote the responsible use of science in public policy;
- Strengthen and diversify the science and technology workforce;
- Foster education in science and technology for everyone;
- Increase public engagement with science and technology; and advance international cooperation in science.


AAAS
ADVANCING SCIENCE, SERVING SOCIETY


halcyon
DIALOGUE