AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

ARTIFICIAL INTELLIGENCE: WITH GREAT POWER COMES GREAT RESPONSIBILITY

JOINT HEARING

BEFORE THE

SUBCOMMITTEE ON RESEARCH AND TECHNOLOGY
&
SUBCOMMITTEE ON ENERGY

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
HOUSE OF REPRESENTATIVES

ONE HUNDRED FIFTEENTH CONGRESS

SECOND SESSION

JUNE 26, 2018

Serial No. 115–67

Printed for the use of the Committee on Science, Space, and Technology

Available via the World Wide Web: http://science.house.gov

U.S. GOVERNMENT PUBLISHING OFFICE
30–877PDF WASHINGTON : 2018

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HON. LAMAR S. SMITH, Texas, *Chair*

FRANK D. LUCAS, Oklahoma
DANA ROHRABACHER, California
MO BROOKS, Alabama
RANDY HULTGREN, Illinois
BILL POSEY, Florida
THOMAS MASSIE, Kentucky
RANDY K. WEBER, Texas
STEPHEN KNIGHT, California
BRIAN BABIN, Texas
BARBARA COMSTOCK, Virginia
BARRY LOUDERMILK, Georgia
RALPH LEE ABRAHAM, Louisiana
GARY PALMER, Alabama
DANIEL WEBSTER, Florida
ANDY BIGGS, Arizona
ROGER W. MARSHALL, Kansas
NEAL P. DUNN, Florida
CLAY HIGGINS, Louisiana
RALPH NORMAN, South Carolina
DEBBIE LESKO, Arizona

EDDIE BERNICE JOHNSON, Texas
ZOE LOFGREN, California
DANIEL LIPINSKI, Illinois
SUZANNE BONAMICI, Oregon
AMI BERA, California
ELIZABETH H. ESTY, Connecticut
MARC A. VEASEY, Texas
DONALD S. BEYER, JR., Virginia
JACKY ROSEN, Nevada
CONOR LAMB, Pennsylvania
JERRY MCNERNEY, California
ED PERLMUTTER, Colorado
PAUL TONKO, New York
BILL FOSTER, Illinois
MARK TAKANO, California
COLLEEN HANABUSA, Hawaii
CHARLIE CRIST, Florida

SUBCOMMITTEE ON RESEARCH AND TECHNOLOGY

HON. BARBARA COMSTOCK, Virginia, *Chair*

FRANK D. LUCAS, Oklahoma
RANDY HULTGREN, Illinois
STEPHEN KNIGHT, California
BARRY LOUDERMILK, Georgia
DANIEL WEBSTER, Florida
ROGER W. MARSHALL, Kansas
DEBBIE LESKO, Arizona
LAMAR S. SMITH, Texas

DANIEL LIPINSKI, Illinois
ELIZABETH H. ESTY, Connecticut
JACKY ROSEN, Nevada
SUZANNE BONAMICI, Oregon
AMI BERA, California
DONALD S. BEYER, JR., Virginia
EDDIE BERNICE JOHNSON, Texas

SUBCOMMITTEE ON ENERGY

HON. RANDY K. WEBER, Texas, *Chair*

DANA ROHRABACHER, California
FRANK D. LUCAS, Oklahoma
MO BROOKS, Alabama
RANDY HULTGREN, Illinois
THOMAS MASSIE, Kentucky
STEPHEN KNIGHT, California
GARY PALMER, Alabama
DANIEL WEBSTER, Florida
NEAL P. DUNN, Florida
RALPH NORMAN, South Carolina
LAMAR S. SMITH, Texas

MARC A. VEASEY, Texas, *Ranking Member*
ZOE LOFGREN, California
DANIEL LIPINSKI, Illinois
JACKY ROSEN, Nevada
JERRY MCNERNEY, California
PAUL TONKO, New York
BILL FOSTER, Illinois
MARK TAKANO, California
EDDIE BERNICE JOHNSON, Texas

CONTENTS

June 26, 2018

	Page
Witness List	2
Hearing Charter	3

Opening Statements

Statement by Representative Barbara Comstock, Chairwoman, Subcommittee on Research and Technology, Committee on Science, Space, and Technology, U.S. House of Representatives	4
Written Statement	6
Statement by Representative Daniel Lipinski, Ranking Member, Subcommittee on Research and Technology, Committee on Science, Space, and Technology, U.S. House of Representatives	8
Written Statement	10
Statement by Representative Lamar Smith, Chairman, Committee on Science, Space, and Technology, U.S. House of Representatives	12
Written Statement	13
Statement by Representative Marc A. Veasey, Ranking Member, Subcommittee on Energy, Committee on Science, Space, and Technology, U.S. House of Representatives	14
Written Statement	15
Statement by Representative Randy K. Weber, Chairman, Subcommittee on Energy, Committee on Science, Space, and Technology, U.S. House of Representatives	16
Written Statement	18
Written statement by Representative Eddie Bernice Johnson, Ranking Member, Committee on Science, Space, and Technology, U.S. House of Representatives	21

Witnesses:

Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office	
Oral Statement	22
Written Statement	25
Mr. Greg Brockman, Co-Founder and Chief Technology Officer, OpenAI	
Oral Statement	40
Written Statement	42
Dr. Fei-Fei Li, Chairperson of the Board and Co-Founder, AI4ALL	
Oral Statement	50
Written Statement	52
Discussion	59

Appendix I: Answers to Post-Hearing Questions

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University	82
Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office	89
Mr. Greg Brockman, Co-Founder and Chief Technology Officer, OpenAI	97
Dr. Fei-Fei Li, Chairperson of the Board and Co-Founder, AI4ALL	105

Page

Appendix II: Additional Material for the Record

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University, written statement 112

Document submitted by Representative Bill Foster, Subcommittee on Research and Technology, Committee on Science, Space, and Technology, U.S. House of Representatives 123

Document submitted by Representative Neal P. Dunn, Subcommittee on Energy, Committee on Science, Space, and Technology, U.S. House of Representatives 150

ARTIFICIAL INTELLIGENCE: WITH GREAT POWER COMES GREAT RESPONSIBILITY

TUESDAY, JUNE 26, 2018

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON RESEARCH AND TECHNOLOGY AND
SUBCOMMITTEE ON ENERGY,
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY,
Washington, D.C.

The Subcommittees met, pursuant to call, at 10:37 a.m., in Room 2318 of the Rayburn House Office Building, Hon. Barbara Comstock [Chairwoman of the Subcommittee on Research and Technology] presiding.

LAMAR S. SMITH, Texas
CHAIRMAN

EDDIE BERNICE JOHNSON, Texas
RANKING MEMBER

Congress of the United States
House of Representatives
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
2321 RAYBURN HOUSE OFFICE BUILDING
WASHINGTON, DC 20515-6301
(202) 225-6371
www.science.house.gov

Artificial Intelligence – With Great Power Comes Great Responsibility

Tuesday, June 26, 2018
10:30 a.m.
2318 Rayburn House Office Building

Witnesses

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University

Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office

Mr. Greg Brockman, Co-Founder and Chief Technology Officer, OpenAI

Dr. Fei-Fei Li, Chairperson of the Board and Co-Founder, AI4ALL

U.S. HOUSE OF REPRESENTATIVES
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HEARING CHARTER

June 26, 2018

TO: Members, Subcommittee on Research and Technology, Subcommittee on Energy

FROM: Majority Staff, Committee on Science, Space, and Technology

SUBJECT: Joint Subcommittee Hearing: "Artificial Intelligence – With Great Power Comes Great Responsibility"

The Subcommittees on Research and Technology and Energy will hold a hearing titled *Artificial Intelligence – With Great Power Comes Great Responsibility*, on Tuesday, June 26, 2018 at 10:30 a.m. in Room 2318 of the Rayburn House Office Building.

Hearing Purpose:

The purpose of the hearing is to understand the state of artificial intelligence technology and the difference between narrow and general intelligence. The hearing will also examine the types of research being conducted to advance artificial general intelligence technology and explore its game-changing potential and implications.

Witness List

- **Dr. Jaime Carbonell**, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University
- **Dr. Tim Persons**, Chief Scientist, U.S. Government Accountability Office
- **Mr. Greg Brockman**, Co-Founder and Chief Technology Officer, OpenAI
- **Dr. Fei-Fei Li**, Chairperson of the Board and Co-Founder, AI4ALL

Staff Contact

For questions related to the hearing, please contact Raj Bharwani of the Majority Staff at 202-225-6371.

Chairwoman COMSTOCK. The Committee on Space, Science, and Technology will come to order. Without objection, the Chair is authorized to declare recesses of the Committee at any time.

Good morning, and welcome to today's hearing entitled "Artificial Intelligence—With Great Power Comes Great Responsibility."

I now recognize myself for five minutes for an opening statement.

First, I would like to note that one of our witnesses, Dr. Jaime Carbonell from Carnegie Mellon University, is unable to be here today due to a medical emergency. We wish him well and a speedy recovery, and, without objection, we'll ensure his written testimony is made part of the hearing record.

[The prepared statement of Mr. Carbonell appears in Appendix II]

Chairwoman COMSTOCK. One of the reasons I've been looking forward to today's hearing is to get a better sense from our witnesses about the nuances of the term artificial intelligence and implications for our society in a future where AI is ubiquitous.

Of course, one might say AI is already pervasive. Since the term was first coined in the 1950s, we have made huge advances in the field of artificial narrow intelligence, which has been applied to many familiar everyday items such as the technology underlying Siri and Alexa.

Called ANI for short, such systems are designed to conduct specific and usually limited tasks. For example, a machine that excels at playing poker wouldn't be able to parallel park a car. Conversely, AGI, or artificial general intelligence, refers to intelligent behavior across a range of cognitive tasks. If you enjoy science fiction movies, this definition may conjure up scenes from any number of classics such as Blade Runner, The Matrix, or The Terminator.

For many individuals, the term AGI invokes images of robots or machines with human intelligence. As it turns out, we are decades away from realizing such AGI systems. Nevertheless, discussions about AGI and a future in which AGI is commonplace lead to some interesting questions worthy of analysis.

For example, Elon Musk has been quoted as saying that AI, quote, "is a fundamental risk to the existence of human civilization" and poses "vastly more risk" than North Korea. Does that mean that AGI may evolve to a point one day when we will lose control over machines of our own creation? As farfetched as that sounds, minds and scientists are certainly discussing such questions.

For the short term, however, my constituents are concerned about less existential issues that usually accompany new and evolving technologies, topics such as cybersecurity, protecting our privacy, and impacts to our nation's economy and to jobs.

I am an original cosponsor of a bill introduced earlier this year titled the AI JOBS Act of 2018 to help our workplace prepare for the ways AI will shape the economy of the future. I will also introduce legislation today to reauthorize the National Institute of Standards and Technology, which includes language directing NIST to support development of artificial intelligence and data science.

There is immense potential for AGI to help humans and to help our economy and all of the issues we're dealing with today, but that potential is also accompanied by some of the concerns that we will discuss today. I look forward to what our panel has to share with us about the bright as well as the challenging sides of the future with AGI.

[The prepared statement of Chairwoman Comstock follows:]

COMMITTEE ON
SCIENCE, SPACE, & TECHNOLOGY
Lamar Smith, Chairman

For Immediate Release
June 26, 2018

Media Contact: Heather Vaughan, Bridget Dunn
(202) 225-6371

Statement by Chairwoman Barbara Comstock (R-Va.)

Artificial Intelligence – With Great Power Comes Great Responsibility

Chairwoman Comstock: I would like to say first that one of our witnesses, Dr. Jaime Carbonell from Carnegie Mellon University, is unable to be here today due to a medical emergency. We wish him well and a speedy recovery, and without objection will ensure his written testimony is made part of the hearing record.

One of the reasons I have been looking forward to today's hearing is to get a better sense from our witnesses about the nuances of the term artificial intelligence (AI), and implications for our society in a future where AI is ubiquitous.

Of course, one might say AI is already pervasive. Since the term was first coined in the 1950s, we have made huge advances in the field of artificial *narrow* intelligence, which has been applied to many familiar every-day items such as the technology underlying Siri and Alexa.

Called ANI for short, such systems are designed to conduct specific and usually limited tasks. For example, a machine that excels at playing poker wouldn't be able to parallel park my car.

Conversely, AGI, or artificial *general* intelligence, refers to intelligent behavior across a range of cognitive tasks. If you enjoy science fiction movies, this definition may conjure up scenes from any number of classics such as "Blade Runner", "The Matrix" or "The Terminator".

For many individuals, the term AGI invokes images of robots or machines with human intelligence. As it turns out, we are decades away from realizing such AGI systems. Nevertheless, discussions about AGI and a future in which AGI is commonplace lead to some interesting questions worthy of analysis.

For example, Elon Musk has been quoted as saying that AI "'is a fundamental risk to the existence of human civilization' and poses 'vastly more risk' than North Korea." Does that mean that AGI may evolve to a point one day when we will lose control over machines of our own creation? As far-fetched as that sounds, great minds are certainly discussing such questions.

For the short term however, my constituents are concerned about less existential issues that usually accompany new or evolving technologies—topics such as cybersecurity, privacy and impacts to our nation's economy and American jobs.

I am an original cosponsor of a bill introduced earlier this year titled the AI JOBS Act of 2018 to help our workforce prepare for the ways AI will shape the economy of the future. I will also introduce legislation today to reauthorize the National Institute of Standards and Technology (NIST), which includes language directing NIST to support development of artificial intelligence and data science.

There is immense potential for AGI to help humans. But that potential is also accompanied by some of the concerns I just referenced. I look forward to what our panel has to share with us about the bright and the dark side of a future with AGI.

###

Chairwoman COMSTOCK. I now recognize the Ranking Member of the Research and Technology Subcommittee, the gentleman from Illinois, Mr. Lipinski, for his opening statement.

Mr. LIPINSKI. Thank you, Chairwoman Comstock, and thank you to Chairman Weber for holding this hearing to understand the current state of artificial intelligence technology.

Because of the rapid development of computational power, the capacity of AI to perform new and more complicated tasks is quickly advancing. Depending on who you ask, AI is the stuff of dreams or nightmares. I believe it is definitely the former, and I strongly fear that it could also be the latter.

The science fiction fantasy worlds depicted on Hollywood's big and small screens alike capture imaginations about what the world might be like if humans and highly intelligent robots shared the Earth. Today's hearing is an opportunity to begin to understand the real issues in AI and to begin to move forward with informed science-based policymaking. This is a hearing that we may remember years from now hopefully as a bright beginning of a new era.

Current AI technologies touch a broad scope of industries and sectors, including manufacturing, transportation, energy, health care, and many others. As we will hear from the witnesses today, artificial intelligence can be classified as artificial general intelligence or artificial narrow intelligence. From my understanding, it is applications of the latter such as machine learning that are underlying technologies that support some of the services and devices widely used by Americans today. These include virtual assistants such as Siri and Alexa, translation services such as Google Translate, and autonomous vehicle technologies. As the capabilities of AI improve, it will undoubtedly become a more essential part of our lives and our economy.

While technology developers and industry look forward to making great strides in AI, I want to make sure my colleagues and I in Congress are asking the tough questions and carefully considering the most crucial roles that the federal government may have in shaping the future of AI. Federal investments in AI research are long-standing, and we must consider the appropriate balance and scope of federal involvement as we begin to better understand the various roles AI will play in our society.

We are not starting from scratch in thinking about the appropriate role of the federal government in this arena. In 2016, the White House issued the National Artificial Intelligence Research and Development Strategic Plan that outlines seven priorities for federally funded AI research. These included making long-term investments in AI, developing effective methods for human AI collaboration, and addressing the ethical, legal, and societal implications of AI, additional issues to address our safety and security, public data sets, standards, and workforce needs.

Earlier this year, the Government Accountability Office issued a technology assessment report led by one of our witnesses, Dr. Persons, titled "Artificial Intelligence: Emerging Opportunities, Challenges, and Implications." While noting significant potential for AI to improve many industries including finance, transportation, and cybersecurity, the report also noted areas where research is still needed, including how to optimally regulate AI, how to ensure the

availability and use of high-quality data, understanding AI's effects on employment and education, and the development of computational ethics to guide the decisions made by software.

These are all critical issues, but more and more I hear concern and widely varying predictions about AI's impact on jobs. AI has the potential to make some job functions safer and more efficient, but it also may replace others. We need to ask what are the long-term projections for the job market as AI grows? In this context, we also need to ask how well do our AI capabilities compare to those of other countries? What education, skills, and retraining will the workforce of the future need? These are very important questions as we think about ensuring a skilled workforce of the future that will help solidify U.S. leadership in AI as other countries vie for dominance in the field. If AI threatens some careers, it likely creates many others. We need to consider what Congress should do to shape this impact to make sure Americans are ready for it and make sure the benefits of AI are distributed widely.

One other obvious issue of major concern when it comes to AI is ethics. There are many places where this becomes relevant. Currently, we need to grapple with issues regarding the data that are being used to educate machines. Biased data will lead to biased results from seemingly objective machines.

A little further down the line are many difficult questions being raised in science fiction about a world of humans and intelligent robots. These are questions we will likely be called on to deal with in Congress, and we need to be ready.

I want to thank all of our witnesses for being here today, and I look forward to your testimony. I'll yield back.

[The prepared statement of Mr. Lipinski follows:]

OPENING STATEMENT
Ranking Member Daniel W. Lipinski (D-IL)
of the Subcommittee on Research and Technology

House Committee on Science, Space, and Technology
Subcommittee on Research and Technology
Subcommittee on Energy
Artificial Intelligence: With Great Power Comes Great Responsibility
June 26, 2018

Thank you Chairwoman Comstock and Chairman Weber for holding this hearing to understand the current state of artificial intelligence technology. Because of the rapid development of computational power, the capacity of AI to perform new and more complicated tasks is quickly advancing.

Depending on who you ask, AI is the stuff dreams or nightmares. I believe it is definitely the former and I strongly fear that it could also be the latter. The science fiction fantasy worlds depicted on Hollywood's big and small screens alike capture our imaginations about what the world might be like if humans and highly intelligent robots shared the Earth. Today's hearing is an opportunity to begin to understand the real issues in AI, and to begin to move forward with informed, science-based policymaking. This is a hearing that we may remember years from now, hopefully as the bright beginning of a new era.

Current AI technologies touch a broad scope of industries and sectors, including manufacturing, transportation, energy, healthcare, and many others. As we will hear from the witnesses today, artificial intelligence can be classified as artificial general intelligence or artificial narrow intelligence. From my understanding, it is applications of the latter, such as machine learning, that are the underlying technologies that support some of the services and devices widely used by Americans today. These include virtual assistants, such as Siri and Alexa, translation services, such as Google Translate, and autonomous vehicle technologies. As the capabilities of AI improve, it will undoubtedly become a more essential part of our lives and our economy.

While technology developers and industry look forward to making great strides in AI, I want to make sure my colleagues and I here in Congress are asking the tough questions and carefully considering the most crucial roles that the federal government may have in shaping the future of AI. Federal investments in AI research are longstanding, and we must consider the appropriate balance and scope of federal involvement as we begin to better understand the various roles AI will play in our society. We are not starting from scratch in thinking about the appropriate role of the federal government in this arena. In 2016, the White House issued the *National Artificial Intelligence Research and Development Strategic Plan* that outlined seven priorities for federally-funded AI research. These included making long-term investments in AI; developing effective methods for human-AI collaboration; and addressing the ethical, legal, and societal implications of AI. Additional issues to address are safety and security, public datasets, standards, and workforce needs.

Earlier this year, the Government Accountability Office issued a technology assessment report, led by one of our witnesses, Dr. Persons, titled *Artificial Intelligence: Emerging Opportunities, Challenges, and Implications.* While noting significant potential for AI to improve many

industries, including finance, transportation, and cybersecurity, the report also noted areas where research is still needed, including how to optimally regulate AI, how to ensure the availability and use of high quality data, understanding AI's effects on employment and education, and the development of computational ethics to guide the decisions made by software.

These are all critical issues, but more and more I hear concern - and widely varying predictions - about AI's impact on jobs. AI has the potential to make some job functions safer and more efficient, but may replace others. We need to ask, what are the long-term projections for the job market as AI grows? In this context we need to ask, how well do our AI capabilities compare to those of other countries? What education, skills, and retraining will the workforce of the future need? These are very important questions as we think about ensuring a skilled workforce for the future that will help solidify U.S. leadership in AI as other countries vie for dominance in this field. If AI threatens some careers, it likely creates many others. We need to consider what Congress should do to shape this impact, make sure Americans are ready for it, and make sure the benefits of AI are distributed widely.

One other obvious issue of major concern when it comes to AI is ethics. There are many places where this becomes relevant. Currently we need to grapple with issues regarding the data that are being used to "educate" machines. Biased data will lead to biased results from seemingly "objective" machines. A little further down the line are many difficult questions being raised in science fiction about a world of humans and intelligent robots. These are questions we will likely be called on to deal with in Congress and we need to be ready.

I want to thank all of our witnesses for being here today and I look forward to hearing their testimony. I yield back.

Chairwoman COMSTOCK. Thank you, Mr. Lipinski.

And I now recognize the Chairman of the Energy Subcommittee, the gentleman from Texas, Mr. Weber, for his opening statement.

Mr. WEBER. Madam Chair, can I defer to the Chairman of the full Committee for his statement?

Chairwoman COMSTOCK. Yes, you may.

Mr. WEBER. Thank you.

Chairman SMITH. Thank you, Madam Chair. Thank you, Mr. Chairman. I didn't know you were going to do that.

Madam Chair, often unknown to us, advances in artificial intelligence, or AI, touch many aspects of our lives. In the area of cybersecurity, AI reduces our reaction times to security threats. In the field of agriculture, AI monitors soil moisture and targets crop watering. And in the transportation lane, AI steers self-driving cars and manages intelligent traffic systems. Multiple technical disciplines, including quantum computing science, converge to form AI.

Tomorrow, the Science Committee will mark up the National Quantum Initiative Act, which establishes a federal program to accelerate quantum research and development. This is a bipartisan bill that Ranking Member Eddie Bernice Johnson and I and others will introduce today. My hope is that every member of the committee will sponsor it or at least a majority.

Transforming our current quantum research into real-world applications will create scientific and technological discoveries, especially in the field of artificial intelligence. These discoveries will stimulate economic growth and improve our global competitiveness, important considerations in light of China's advances in artificial intelligence and quantum computing. By some accounts, China is investing $7 billion in AI through 2030, and $10 billion in quantum research.

The European Union has also issued a preliminary plan outlining a $24 billion public-private investment in AI between 2018 and 2020. And Russian President Putin has noted that, quote, "The leader in AI will rule the world," end quote. No doubt that's appealing to him. Yet, the Department of Defense's unclassified investment in AI was only $600 million in 2016, while federal spending on quantum totals only about $250 million a year.

The Committee will mark up a second piece of legislation to reauthorize the National Institute of Standards and Technology. The bill directs NIST to continue supporting the development of artificial intelligence and data science, including the development of machine learning and other artificial intelligence applications. It is simply vital to our nation's future that we accelerate our quantum computing and artificial intelligence efforts.

Thank you, Madam Chair, and I yield back.

[The prepared statement of Chairman Smith follows:]

COMMITTEE ON
SCIENCE, SPACE, & TECHNOLOGY
Lamar Smith, Chairman

For Immediate Release
June 26, 2018

Media Contacts: Heather Vaughan, Bridget Dunn
(202) 225-6371

Statement by Chairman Lamar Smith (R-Texas)

Artificial Intelligence – With Great Power Comes Great Responsibility

Chairman Smith: Often unknown to us, advances in artificial intelligence—or AI—touch many aspects of our lives. In the area of cybersecurity, AI reduces our reaction times to security threats. In the field of agriculture, AI monitors soil moisture and targets crop watering. And in the transportation lane, AI steers self-driving cars and manages intelligent traffic systems. Multiple technical disciplines including quantum computing science converge to form AI.

Tomorrow, the Science Committee will mark up the National Quantum Initiative Act, which establishes a federal program to accelerate quantum research and development. This is a bipartisan bill that Ranking Member Eddie Bernice Johnson and I and others will introduce today. My hope is that every member of the committee will sponsor it.

Transforming our current quantum research into real world applications will create scientific and technological discoveries, especially in the field of artificial intelligence.

These discoveries will stimulate economic growth and improve our global competitiveness—important considerations in light of China's advances in artificial intelligence and quantum computing. By some accounts, China is investing $7 billion in AI through 2030, and $10 billion in quantum research.

The European Union also has issued a preliminary plan outlining a $24 billion public-private investment in AI between 2018 and 2020. And Russian President Vladimir Putin has noted that "the leader in AI will 'rule the world.'" No doubt that's appealing to him.

Yet, the Department of Defense's unclassified investment in AI was only $600 million in 2016 while federal spending on quantum totals about $250 million a year.

The committee will mark-up a second piece of legislation to reauthorize the National Institute of Standards and Technology (NIST). The bill directs NIST to continue supporting the development of artificial intelligence and data science including the development of machine learning and other artificial intelligence applications.

It is vital to our nation's future that we accelerate our quantum computing and AI efforts.

###

Chairwoman COMSTOCK. Thank you. And I now recognize the Ranking Member of the Energy Subcommittee, the gentleman from Texas, Mr. Veasey, for an opening statement.

Mr. VEASEY. I want to thank you, Chairwoman Comstock and Chairman Weber, for holding this hearing today, and thank you for all of the witnesses for providing expertise on this topic. I'm looking forward to hearing what everyone has to say today.

America, of course, is a country of innovation, and in the digital world of today, more and more industries are relying on advanced technologies and connectivity to overcome new challenges. Artificial intelligence and big data are impacting every facet of production and commerce. AI has the ability to mimic cognitive functions such as problem-solving and learning, making it a critical resource as we encounter never-before-seen problems. Those in the energy sector have already seen improvements in productivity and efficiency and can expect to see even more advancement in the coming years.

AI can be used to process and analyze data in previously unexplored ways. Technology such as sensor-equipped aircraft engines, locomotive, gas, and wind turbines are now able to track production efficiency and wear and tear on vital machinery.

AI could also significantly improve our ability to detect failures before they occur and prevent disasters, saving money, time, and lives. And through the use of analytics, sensors, and operational data, AI can be used to manage, maintain, and optimize systems ranging from energy storage components to power plants to the electric grid. As digital technologies revolutionize the energy sector, we must ensure safe and responsible execution of these processes.

AI systems can learn and adapt through continuous modeling of interaction and data feedback. Production must be put in place to guarantee the integrity of these mechanisms as they evaluate mass quantities of machine and user data. With Americans' right to privacy under threat, security of these connected systems is of the utmost importance.

Nevertheless, I'm excited to learn about the valuable benefits that AI may be able to provide for our economy and our well-being alike. With a Gartner research study reporting that AI will generate 2.3 million jobs by 2020, that's a lot of jobs. The growth AI will bring not only to the energy sector but to health care, transportation, education, and so many others will help ensure the prosperity of our nation.

I look forward to seeing what light our witnesses can shed on these topics and what we can do in Congress to help enable the development and deployment of these promising technologies.

Madam Chairwoman, I yield back the balance of my time.

[The prepared statement of Mr. Veasey follows:]

OPENING STATEMENT
Ranking Member Marc Veasey (D-TX)
of the Subcommittee on Energy

House Committee on Science, Space, and Technology
Subcommittee on Research and Technology
Subcommittee on Energy
Artificial Intelligence – With Great Power Comes Great Responsibility
June 26, 2018

Thank you, Chairwoman Comstock and Chairman Weber for holding this hearing today, and thank you to all our witnesses for providing their expertise on this topic. I am looking forward to hearing what you all have to say.

America is a country of innovation, and in the digital world of today more and more industries are relying on advanced technologies and connectivity to overcome new challenges. Artificial Intelligence and Big Data are impacting every facet of production and commerce. AI has the ability to mimic cognitive functions such as problem solving and learning, making it a critical resource as we encounter never-before-seen problems.

Those in the energy sector have already seen improvements in productivity and efficiency and can expect to see even more advancements in the coming years. AI can be used to process and analyze data in previously unexplored ways. Technologies such as sensor-equipped aircraft engines, locomotives, gas turbines and wind turbines are now able to track production efficiency and the wear and tear on vital machinery. With that technology, we can expect significant reductions in fuel consumption as well as carbon emissions.

AI could also significantly improve our ability to detect failures before they occur and prevent disasters, saving money, time, and lives. And through the use of analytics, sensors, and operational data, AI can be used to manage, maintain, and optimize systems ranging from energy storage components to power plants to the electric grid.

As digital technologies revolutionize the energy sector, we must ensure safe and responsible execution of these processes. AI systems learn and adapt through continuous modelling of interaction data and feedback. Precautions must be put in place to guarantee the integrity of these mechanisms as they evaluate mass quantities of machine and user data. With Americans right to privacy under threat, security of these connected systems is of the utmost importance.

Nevertheless, I am excited to learn more about valuable benefits that AI may be able to provide for our economy and wellbeing alike. With a Gartner research study reporting that AI will generate 2.3 million jobs by 2020, the growth AI will bring not only to the energy sector, but to healthcare, transportation, education, and more, will help ensure the prosperity of our nation.

I look forward to seeing what light our witnesses can shed on these topics, and what we in Congress can do to enable the development and deployment of these promising technologies. Thank you, and I yield back the remainder of my time.

Chairwoman COMSTOCK. Thank you. And I now recognize Mr. Weber for his opening statement.

Mr. WEBER. Thank you, Madam Chair.

Today, we will hear from a panel of experts on next-generation artificial intelligence, AI as we've all heard it described. And while some have raised concerns about the negative consequences of AI, this technology has the potential to solve fundamental science problems and improve everyday life. In fact, it's likely that everyone in this room benefits from artificial intelligence. For example, users of voice assistants, online purchase prediction, fraud detection that the gentleman from Texas mentioned, and music recommendation services are already utilizing aspects of this technology in their day-to-day life.

In the past few years, the use of AI technology has rapidly expanded due to the increase in the volume of data worldwide, and to the proliferation of advanced computing hardware that allows for the powerful parallel processing of this data. The field of AI has broadened to include other advanced computing disciplines such as machine learning. We've heard about neural networks, deep learning computer vision, and natural language processing, just to name a few. These learning techniques are key to the development of AI technologies and can be used to explore complex relationships and produce previously unseen results on unprecedented timescales.

The Department of Energy, DOE, is the nation's largest federal supporter of basic research in the physical science, with expertise in big-data science, high-performance computing, advanced algorithms, and data analytics and is uniquely positioned to enable fundamental research in AI and machine learning.

DOE's Office of Science Advanced Scientific Computing Research program, or ASCR as we call it, program develops next-generation supercomputing systems that can achieve the computational power needed for this type of critical research. This includes the Department's newest and most powerful supercomputer called Summit, which just yesterday, just yesterday was ranked as the fastest computing system in the entire world.

AI also has broad applications in the DOE mission space. In materials science, AI helps researchers speed the experimental process and discover new compounds faster than ever before. In high-energy physics, AI finds patterns in atomic and particle collisions previously unseen by scientists.

In fusion energy research, AI modeling predicts plasma behavior that will assist in building tokamak reactors, making the best of our investments in space. Even in fossil fuel energy production, AI systems will optimize efficiency and predict needed maintenance at power-generating facilities. AI technology has the potential to improve computational science methods for any big-data problem, any big-data problem. And with the next generation of supercomputers, the exascale computing systems that DOE is expected to field by 2021, American researchers utilizing AI technology will be able to track even bigger challenges.

We cannot afford to fall behind in this compelling area of research, and big investments in AI by China and Europe already threaten U.S. dominance in this field. With the immense potential

for AI technology to answer fundamental scientific challenges, it's quite clear we should prioritize this research.

We should maintain, I will add, American competitive edge and American exceptionalism. This will help us to do that.

I want to thank our accomplished panel of witnesses for their testimony today, and I look forward to hearing what role Congress can play and should play in advancing this critical area of discovery science.

And, Madam Chair, I yield back.

[The prepared statement of Mr. Weber follows:]

COMMITTEE ON
SCIENCE, SPACE, & TECHNOLOGY
Lamar Smith, Chairman

For Immediate Release
June 26, 2018

Media Contacts: Heather Vaughan, Bridget Dunn
(202) 225-6371

Statement by Chairman Randy Weber (R-Texas)

Artificial Intelligence – With Great Power Comes Great Responsibility

Chairman Weber: Today, we will hear from a panel of experts on next generation artificial intelligence—or AI—technologies. While some have raised concerns about the negative consequences of AI, this technology has the potential to solve fundamental science problems and improve everyday life.

In fact, it is likely that everyone in this room benefits from artificial intelligence. For example, users of voice assistants, online purchase prediction, fraud detection, and music recommendation services are already utilizing aspects of this technology in their day to day life.

In the past few years, the use of AI technology has rapidly expanded due to the increase in the volume of data worldwide, and to the proliferation of advanced computing hardware that allows for the powerful parallel processing of this data. The field of AI has broadened to include other advanced computing disciplines such as machine learning, neural networks, deep learning computer vision and natural language processing, to name a few.

These learning techniques are key to the development of AI technologies, and can be used to explore complex relationships and produce previously unseen results on unprecedented timescales.

The Department of Energy (DOE), is the nation's largest federal supporter of basic research in the physical sciences, with expertise in big data science, high performance computing, advanced algorithms and data analytics and is uniquely positioned to enable fundamental research in AI and machine learning.

DOE's Office of Science Advanced Scientific Computing Research program, or ASCR, program develops next generation supercomputing systems that can achieve the computational power needed for this critical research. This includes the Department's newest and most powerful supercomputer, called Summit, which just yesterday was ranked as the fastest computing system in the world.

AI also has broad applications in the DOE mission space. In materials science, AI helps researchers speed the experimental process and discover new compounds faster than ever before. In high energy physics, AI finds patterns in atomic and particle collisions previously unseen by scientists.

In fusion energy research, AI modeling predicts plasma behavior that will assist with building tokamak reactors, making the best of our investment in this space. Even in fossil energy production, AI systems will optimize efficiency and predict needed maintenance at power generating facilities.

AI technology has the potential to improve computational science methods for any big data problem.

And with the next generation of supercomputers, the exascale computing systems that DOE is expected to field by 2021, American researchers utilizing AI technology will be able to tackle even bigger challenges.

We can't afford to fall behind in this compelling area of research—and big investments in AI by China and Europe already threaten U.S. dominance in this field. With the immense potential for AI technology to answer fundamental scientific challenges, it's clear we should prioritize this research.

I want to thank our accomplished panel of witnesses for their testimony today and I look forward to hearing what role Congress should play in advancing this critical area of discovery science.

###

[The prepared statement of Full Committee Ranking Member Eddie Bernice Johnson]

OPENING STATEMENT

Ranking Member Eddie Bernice Johnson (D-TX)

House Committee on Science, Space, and Technology
Subcommittee on Research and Technology
Subcommittee on Energy
Artificial Intelligence: With Great Power Comes Great Responsibility
June 26, 2018

Thank you Chairwoman Comstock, Ranking Member Lipinski, Chairman Weber, and Ranking Member Veasey for holding this important hearing, and thank you to the witnesses for being here today.

Artificial intelligence, or AI, is already affecting the way we make decisions and interact in our world. When fully developed, AI will likely have a large impact across all sectors, industries, and demographics. This committee should support policies that ensure safe, responsible development of AI that will benefit society and make the world a better place. The Science committee should also be stewards for bridging the digital divide and supporting efforts to include broad participation in AI and to prevent algorithmic bias. AI's current and potential applications are numerous, including virtual personal assistants and predictive technologies that enhance customer interactions.

From healthcare to national security and everything in-between, advancement of AI depends on building foundational knowledge in the many fields that contribute to AI, including the social, behavioral, and economic sciences. Federal investment in basic research for AI is critical. The top spot for world leadership is open for the taking. Our counterparts across the globe have already staked their claim and are investing heavily in R&D and in producing a well-qualified talent pool.

Strategic, robust federal investments in AI R&D and in educating and training our future AI workforce will be critical if we are to remain a leader in AI. These efforts should be directed toward attracting students to the STEM fields and to ensuring our seasoned workforce gains the skills necessary to thrive in an AI-driven economy.

This committee has been at the forefront of supporting game-changing research that advances the boundaries of science and creates completely new industries that could not have been imagined by earlier generations. We have the opportunity to do that with AI if we commit to sustainable funding of research, education, and coordination activities.

I look forward to hearing more about what the future holds for this exciting technology.

Thank you and I yield back.

Chairwoman COMSTOCK. Thank you. And I will now introduce today's witnesses. Our first witness today is Dr. Tim Persons, Chief Scientist at the U.S. Government Accountability Office. He also serves as a Director for GAO's Center for Science, Technology, and Engineering. Dr. Persons received a Bachelor of Science in physics from James Madison University and a Master of Science in nuclear physics from Emory University. He also earned a Master of Science in computer science and Ph.D. in biomedical engineering, both from Wake Forest University.

Next, we have Mr. Greg Brockman, our second witness, who is Cofounder and Chief Technology Officer of OpenAI, a nonprofit artificial intelligence research company. Mr. Brockman is an investor in over 30 startups and a board member of the Stellar digital currency system. He was previously the CTO of Stripe, a payments startup now valued at over $9 billion. And he studied mathematics at Harvard and computer science at MIT.

And our final witness is Dr. Fei-Fei Li, Chairperson of the Board and Cofounder of AI4ALL. In addition, Dr. Li is a Professor in the Computer Science Department at Stanford and the Director of the Stanford Artificial Intelligence Lab. In 2017, Dr. Li also joined Google Cloud as Chief Scientist of AI and machine learning. Dr. Li received her Bachelor of Arts in physics from Princeton and her Ph.D. in electrical engineering from the California Institute of Technology.

I now recognize Dr. Persons for five minutes to present his testimony.

TESTIMONY OF DR. TIM PERSONS, CHIEF SCIENTIST, U.S. GOVERNMENT ACCOUNTABILITY OFFICE

Dr. PERSONS. Good morning. Thank you, Chairwoman Comstock, Chairman Weber, Ranking Members Lipinski and Veasey and Members of the Subcommittee. I'm pleased to be here today to discuss GAO's technology assessment on artificial intelligence. To ensure the United States remains a leader in AI innovation, special attention will be needed for our education and training systems, regulatory structures, frameworks for privacy and civil liberties, and our understanding of risk management in general.

AI holds substantial promise for improving human life, increasing the nation's economic competitiveness, and solving some of society's most pressing challenges. Yet, as a disruptive technology, AI poses risks that could have far-reaching effects on, for example, the future labor force, economic inclusion, and privacy and civil liberties, among others.

Today, I'll summarize three key insights arising from our recent work. First, the distinction between narrow versus general AI; second, the expected impact of AI on jobs, competitiveness, and workforce training; and third, the role the federal government can play in research, standards development, new regulatory approaches, and education.

Regarding narrow versus general AI, narrow AI refers to applications that are task-specific such as tax preparation software, voice and face recognition systems, and algorithms that classify the content of images. General AI refers to a system exhibiting intelligence

on par with or possibly exceeding that of humans. While science fiction has helped general AI capture our collective imaginations for some time, it is unlikely to be fully achieved for decades if at all. Even so, considerable progress has been made in developing narrow AI applications that outperform humans in specific tasks and are thus invoking crucially important economic policy and research considerations.

Regarding jobs, competition, and the workforce, there is considerable uncertainty about the extent to which jobs will be displaced by AI and how many—how much any losses will be offset by job creation. In the near term, displacement to certain jobs such as call-center or retail workers may be particularly vulnerable to automation. However, in the long term, demand for skills that are complementary to AI is expected to increase, resulting in greater productivity. To better understand the impact of AI on employment moving forward, several experts underscored the need for new data and methods to enable greater insight into this issue.

Regarding the role of the federal government, it will continue its crucial role in research and data-sharing, contributions to standards development, regulatory approaches, and education. One important research area of the federal government could support is enhancing the explainability of AI, which could help establish trust in the behavior of AI systems. The federal government could also incentivize data-sharing, including federal data that are subject to limitations for how they can be used, as well as creating frameworks for sharing data to improve the safety and security of AI systems. Such efforts may include supporting standards for explainability; data labeling and safety, including risk assessment; and benchmarking of AI performance against the status quo. It's always risk versus risk.

Related to this, new regulatory approaches are needed, including the development of regulatory sandboxes for testing AI products, services, and business models, especially in industries like transportation, financial services, and health care. GAO's recent report on fintech found, for example, that regulators use sandboxes to gain insight into key questions, issues, and unexpected risks that may arise out of the emerging technologies. New rules governing intellectual property and data privacy may also be needed to manage the deployment of AI.

Finally, education and training will need to be reimagined so workers have the skills needed to work with and alongside emerging AI technologies. For the United States to remain competitive globally and effectively manage AI systems, its workers will need a deeper understanding of probability and statistics across most if not all academic disciplines, that is, not just the physical, engineering, and biological sciences, as well as competency and ethics, algorithmic auditability, and risk management.

In conclusion, the emergence of what some have called the fourth industrial revolution and AI's key role in driving it will require new frameworks for business models and value propositions for the public and private sectors alike. Even if AI technologies were to cease advancing today, no part of society or the economy would be directly or indirectly untouched by its transformative effects.

I thank the committee leadership of the committees. Thanks to the members here for your holding a hearing on this very important topic today for such a time as this.

Madam Chairwoman, Mr. Chairman, Ranking Members, this concludes my prepared remarks. I would be happy to respond to any questions that you or other Members of the Subcommittees have at this time.

[The prepared statement of Dr. Persons follows:]

United States Government Accountability Office

GAO

Testimony

Before the Subcommittees on Research and Technology and Energy, Committee on Science, Space, and Technology, House of Representatives

For Release on Delivery
Expected at 10:30 a.m. ET
Tuesday, June 26, 2018

ARTIFICIAL INTELLIGENCE

Emerging Opportunities, Challenges, and Implications for Policy and Research

Statement of Timothy M. Persons, Chief Scientist
Applied Research and Methods

GAO-18-644T

Chairwoman Comstock, Chairman Weber, and Ranking Members Lipinski and Veasey:

Thank you for the opportunity to discuss our work on artificial intelligence (AI). My testimony today summarizes our March 2018 technology assessment entitled *Artificial Intelligence: Emerging Opportunities, Challenges, and Implications.*[1]

According to experts, AI holds substantial promise for not only improving human life and economic competitiveness in a variety of capacities, but also helping to solve some of society's most pressing challenges. At the same time, AI poses new risks and has the potential to displace workers in some sectors, requires new skills and adaptability to changing workforce needs, and could exacerbate socioeconomic inequality.

Our March 2018 report and my statement today address the following topics:

- How has AI evolved over time?
- According to experts, what are the opportunities and future promise, as well as the principal challenges and risks, of AI?
- According to experts, what are the policy implications and research priorities resulting from advances in AI?

For our March 2018 report, the Comptroller General of the United States convened a Forum on Artificial Intelligence, a meeting of 21 expert participants held on July 6 and 7, 2017, with the assistance of the National Academy of Sciences.[2] The work for the report also included a review of relevant literature and consultation with additional subject-matter experts. Additional information about our scope and methodology can be found in our report. We performed the work on which this testimony is based in accordance with all sections of GAO's Quality Assurance Framework that are relevant to technology assessments.

GAO currently has work underway on how automation is affecting labor markets, which we expect to publish in early 2019. Because of the

[1]GAO, *Artificial Intelligence: Emerging Opportunities, Challenges, and Implications*, GAO-18-142SP (Washington, D.C.: March 28, 2018).

[2]Forum participants were from academia, business, government, and nonprofit organizations. For a complete list of participants, see Appendix II of GAO-18-142SP.

strategic importance of AI in the health care sector, GAO is planning a forum to develop a technology assessment of this area, as well.

The Evolution and Characteristics of AI

Several Definitions and Taxonomies of AI Exist

The field of artificial intelligence can be traced back to a 1956 workshop organized by John McCarthy, held at Dartmouth College. The workshop's goal was to explore how machines could be used to simulate human intelligence. Numerous factors, primarily the trends underlying big data (i.e., increased data availability, storage, and processing power), have contributed to rapid innovation and accomplishments in AI in recent years.[3]

As we noted in our March 2018 technology assessment, there is no single universally accepted definition of AI, but rather differing definitions and taxonomies. In addition to defining AI overall, researchers have distinguished between narrow and general AI. Narrow AI refers to applications that provide domain-specific expertise or task completion, whereas general AI refers to an AI application that exhibits intelligence comparable to a human, or beyond, across the range of contexts in which humans interact. While there has been considerable progress in developing AI that outperforms humans in specific domains, some observers believe that general AI is unlikely to be achieved for decades in the future.

AI Has Been Conceptualized as Having Three Waves of Development

In our March 2018 work, we noted that rather than focusing on a specific definition of AI, it can be understood in terms of the waves in which the technology has developed. Launchbury (2016) provides a framework that conceptualizes AI as having three waves based on differences in capabilities with respect to perceiving, learning, abstracting, and reasoning.[4]

[3]For more on trends underlying big data, see, for example, GAO, *Highlights of a Forum: Data and Analytics Innovation: Emerging Opportunities and Challenges*, GAO-16-659SP (Washington, D.C.: Sept. 20, 2016).

[4]John Launchbury, *A DARPA Perspective on Artificial Intelligence*, 2016.

- **The first wave of AI** is represented by expert knowledge or criteria developed in law or other authoritative sources and encoded into a computer algorithm, which is referred to as an expert system. Examples of expert systems include programs that schedule logistics or prepare taxes.
- **Second-wave AI technology** is based on machine learning, or statistical learning, and includes natural-language processing (e.g., voice recognition) and computer-vision technologies, among others. In contrast to first-wave systems, second-wave systems are designed to perceive and learn. Examples of second-wave systems include voice-activated digital assistants, applications that assist healthcare workers in selecting appropriate treatment options or making diagnoses, and self-driving automated vehicles.
- **Third-wave AI technologies** combine the strengths of first- and second-wave AI and are also capable of contextual sophistication, abstraction, and explanation. An example of third-wave AI is a ship that can navigate the sea without human intervention for a few months at a time while sensing other ships, navigating sea lanes, and carrying out necessary tasks.

As described by Launchbury, we are just at the beginning of the third wave of AI, and further research remains before third-wave technologies become prevalent. An important part of third-wave AI will be developing systems that are not only capable of adapting to new situations, but also are able to explain to users the reasoning behind these decisions.

Forum Participants Identified Several Benefits of Artificial Intelligence and Challenges to Its Development

The increased adoption of artificial intelligence will bring with it several benefits, as well as a number of challenges. According to participants at the forum we convened for our March 2018 technology assessment, both benefits and challenges will need to be carefully considered alongside one another. Figure 1 summarizes selected questions, benefits, and challenges regarding the use of AI in four high-consequence sectors. Participants also stressed that there may be benefits related to AI that cannot yet be predicted or may even be hard to imagine.

Figure 1: Selected Questions Regarding the Use of Artificial Intelligence (AI) in Four High-Consequence Sectors

	Selected Questions
Cybersecurity AI applications face threats from cybersecurity attacks, but AI also may be used as a tool for detecting and defending against attacks.	▸ How can autonomous systems be made secure, without stifling innovation? ▸ How useful is a risk-based approach to determining if machine-learning algorithms adhere to legal requirements or ethical norms?
Automated Vehicles Automated vehicles hold promise for increasing driving safety and providing enhanced mobility, but pose challenges for assuring increased safety.	▸ What is the appropriate regulatory framework for automated vehicle safety assurance? ▸ What are the roles of federal, state, and local governments in infrastructure adaptation and addressing issues of liability and enforcement?
Criminal Justice The use of AI in criminal justice may improve the allocation of law enforcement resources and has the potential to reduce crime and jail populations, but also raises concerns about privacy and civil rights violations.	▸ What are the options for assessing accuracy and the potential for bias in AI data and algorithms? ▸ What are solutions for safeguarding privacy in the collection and use of personal information by AI systems?
Financial Services The use of AI in financial services could improve client services and enhance surveillance monitoring, but also poses challenges to ensuring fair lending, attracting and retaining staff with requisite skills, and maintaining hardware and software.	▸ What are the mechanisms to address ethical considerations, tradeoffs, and protections? ▸ How can regulatory sandboxes be used to test new AI products, services, and business models?

Source: GAO Forum on Artificial Intelligence. | GAO-18-644T

Benefits Identified by Forum Participants

Improved economic outcomes and increased levels of productivity. It may be difficult to accurately predict what AI's impact on the economy could be, according to one forum participant. In previous periods, large investments in automation have been highly correlated with improvements in productivity and economic outcomes, which, according to one forum participant, has led some to believe that transformations as a result of AI could have the same outcome. This same participant noted, however, that no one collects the data needed to measure the impact AI

or other types of advanced automation may have on the economy. According to another participant, whatever the effect that AI will have on productivity in particular, and the economy in general, the changes will occur quickly and be difficult to predict.

Improved or augmented human decision making. AI can be used to gather an enormous amount of data and information from multiple locations, characterize the normal operation of a system, and detect abnormalities much faster than humans can. In addition, AI could be used to create data-informed policy that may help prevent inappropriate or harmful human bias—be it from political pressure or other factors—from creating undesirable results, according to one participant. However, as another participant at the forum noted, AI is no guarantee of freedom from bias. The participant stressed specifically that if the data being used by AI are biased, the results will be biased as well. AI can help prevent inappropriate or harmful human bias, according to this same participant, if it is carefully used, if the assumptions of the models are thoughtfully considered, and, most importantly, if the outputs of the model are constantly and closely verified.

Insights into complex and pressing problems. Some of the participants at our forum believed that AI has the potential to provide insights into—and even help solve—some of the world's most complex and pressing problems. For example, one participant stated that as the number of elderly Americans continues to grow, AI could be used to provide medication management, mobility support, housework, meal preparation, and rehabilitation services to a growing number of people who need assistance with day-to-day activities. In addition, there are other complex and pressing problems that may eventually be solved by the adoption of AI. According to one participant, AI could eventually be used to assure regulatory compliance in the financial sector without unnecessary burden on those being regulated.

Challenges Identified by Forum Participants

Barriers to collecting and sharing data. While not all applications of AI require massive amounts of data, certain applications that use machine learning algorithms do.[5] This can be a problem in sectors where data are not easily aggregated or interpreted or readily available. Such is the case with criminal justice, where the ways in which data are collected and organized vary from jurisdiction to jurisdiction. Such is also true with most vulnerable populations and developing countries, where data have not yet been collected.

Lack of access to adequate computing resources and requisite human capital. Forum participants told us that AI researchers and developers need access to storage and processing, both of which are expensive and sometimes difficult to access at the necessary scale. Some forum participants also shared concerns that the accelerated pace of change associated with AI is straining the education and workforce systems' capacity to train and hire individuals with the appropriate skill sets, leaving many companies struggling to find workers with relevant knowledge, skills, and training.

Adequacy of current laws and regulations. The widespread adoption of AI may, according to some forum participants, have implications regarding the adequacy of current laws and regulations. For example, one participant noted that current patent and copyright laws provide only limited protection for software and business methods and questioned whether these laws will protect the products created by AI. At the same time, one of the participants at the forum raised concerns about ways in which AI could be used to violate civil rights. This participant cautioned, for example, that if law enforcement considers race, class, or gender in AI that is used to assess risk, there is the possibility that a defendant's equal protection rights under the 14th Amendment may be violated, as well as their due process rights under the 5th and 14th Amendments.

Ethical Framework for and Explainability and Acceptance of AI. The adoption of AI also introduces ethical implications. According to a forum participant, there is a need for a system of computational ethics to help AI choose options that reflect agreed-upon values. Moreover, some of the participants at the forum noted that before humans will understand,

[5]Scott W. Bauguess, "The Role of Big Data, Machine Learning, and AI in Assessing Risks: a Regulatory Perspective," U.S. Securities and Exchange Commission, Keynote Address to OpRisk North America 2017, New York, New York, June 21, 2017.

appropriately trust, and be able to effectively manage AI, an AI application or system needs to explain why it took certain actions and why it valued certain variables more than others.

Forum Participants Identified Several Cross-Cutting Policy Considerations Related to AI and Several Areas Where More Research Is Needed

After discussing the benefits and challenges associated with AI, the participants at the forum we convened for our March 2018 technology assessment highlighted a number of policy considerations and areas of future research (see fig. 2).

Figure 2: Implications of Artificial Intelligence (AI) for Policy and Research

Source: GAO Forum on Artificial Intelligence. | GAO-18-644T

Incentivizing data sharing. Forum participants emphasized the need for establishing a "safe space" to protect sensitive information (e.g., intellectual property and brand information) while sharing data. Another participant cautioned that for such a safe space to succeed, it will need to start with a few manufacturers and clearly define the data that are needed and the specific scenarios in which the data will be used.

Certain forum participants also expressed concerns that many potentially useful data are guarded by federal agencies that do not provide access to researchers. Participants noted successful data-sharing efforts through entities such as MITRE and the National Institute of Standards and Technology (NIST). In particular, some participants highlighted data-sharing efforts to improve safety outcomes. For instance, one participant mentioned that researchers at MITRE had credited data-sharing efforts in the aviation industry (employing a safe space) with reducing the number of accidents. Another participant emphasized the importance of sharing data to better understand safety outcomes associated with automated vehicles, stating, "[i]f we're going to trust that these vehicles can go out on the road, we need to verify that, in fact, out on the road, they are as safe as we think they are."

Forum participants highlighted other proposed future data-sharing efforts, citing the benefits of assessing data from multiple sources to improve outcomes. According to one forum participant, the National Science and Technology Council Subcommittee on Machine Learning and Artificial Intelligence is working collaboratively among federal departments and agencies to promote the sharing of government data to help develop innovative solutions for social good. This sharing may include creating training environments—safe spaces—in which sensitive data are protected, among other things.

Another participant noted that in the criminal-justice sector, the federal system could be used as a test bed for various reforms—including data sharing reforms—because the federal system is unified. This participant argued that if the federal system could find a way to share data related to risk assessments and other areas and show that the data are being utilized in an evenhanded way, the reforms pioneered by the federal system would likely migrate down to the individual state systems. This same participant also stated that the Bureau of Justice Assistance and the Bureau of Justice Statistics may be the best positioned to initiate any nationwide data standardization and collection projects.

Improving safety and security. Participants highlighted challenges and opportunities to enhancing the safety and security of system applications from cyber attacks, including those with AI features. One participant said that the costs of cybersecurity in all forms of network computing are not being shared appropriately and that security breaches are much costlier than the security measures that are needed to prevent breaches. This participant said that policymakers will need to consider creating some kind of framework that ensures costs—and liabilities—are appropriately

shared between manufacturers and users. In addition, two participants said that policymakers should consider creating a new regulatory structure to better ensure the safety of automated vehicles.

Updating the regulatory approach. The widespread adoption of AI will have implications for regulators, and lawmakers will need to consider policy options to address these issues, according to multiple forum participants. One participant reinforced the need for regulators to be proactive, including a commitment of resources, because change is occurring so rapidly and in unanticipated ways. For example, as a policy matter going forward, one participant explained, a new regulatory structure for automated vehicles needs to evolve and that, accordingly, the federal government should avoid setting standards prematurely. Another interrelated issue raised by a participant about automated vehicles concerned how liability would be regulated. Currently, according to this participant, the manufacturer of the automated vehicle bears all responsibility for crashes, even if these vehicles improve overall public safety. Some of the participants at the forum also raised concerns about privacy, including ways in which AI could be used by law-enforcement agencies to violate civil liberties, and said that this is an area that needs policy solutions.

In addition, one of the forum participants said that policymakers should consider allowing financial regulators to explore alternative regulatory approaches and reporting mechanisms, leveraging technology to improve and reduce the burden of regulation. In this regard, one participant discussed the merits of "regtech," that is, linking regulation with technology.

Another participant noted that other laws and regulations may need to be adapted to account for the fact that humans may not always be behind decisions that are made by automated systems. For example, this participant discussed laws where intent plays a key role, as is the case in financial market manipulation. If someone programs AI to make money, and it does so in a nefarious way, it is not clear how current laws could be used to prosecute the creator of the AI.

Assessing acceptable risks and ethical decision making. Policymakers need to decide how they are going to measure, or benchmark, the performance of AI and assess the trade-offs, according to one participant, who stressed that the "baseline" is current practice, not perfection (i.e., how humans are performing now, absent AI). As this participant emphasized, "[i]f we have to benchmark [AI] against

perfection...the perfect will be the enemy of the good and we get nowhere." Several participants at the forum emphasized that such regulatory questions should be resolved by a variety of stakeholders, including economists, legal scholars, philosophers, and others involved in policy formulation and decision making, and not solely scientists and statisticians.

Participants at our AI forum also highlighted several areas they believe deserve more research in terms of new regulatory frameworks, data labeling, employment and education, and explainable AI and computational ethics.

Establishing regulatory sandboxes. In finance there is a worldwide movement to create so-called regulatory sandboxes, according to one participant, where regulators can begin experimenting on a small scale and empirically testing new ideas. As this participant explained, regulatory sandboxes would provide a safe haven to assess the results of alternative regulatory approaches.

Developing high-quality labeled data. One participant emphasized the importance of data collection and how to obtain high-quality labeled data. This encompasses improving the quality of the data during data collection. Another participant we spoke with highlighted the merits of developing adequately labeled data sets. As data become more comprehensive and organized, or labeled, in a manner that facilitates machine learning, AI tools can produce more accurate outcomes.

Understanding AI's effect on employment and reimagining training and education. Some forum participants offered mixed views concerning the impacts associated with AI on employment, while acknowledging the uncertainties. For instance, some forum participants noted that job losses in some areas were likely, while noting the potential for job increases in other areas. One participant advocated for research to better understand how jobs have been changing. There is currently no comprehensive federal data source with information on the employment effects AI may have in manufacturing and other segments of the economy. Further, according to two participants, in the absence of a comprehensive data-collection effort, it is unclear which jobs will be created by AI, which jobs may be augmented, or which jobs are likely to be displaced by AI. The widespread adoption of AI also brings with it a need to reevaluate and reimagine training and education, according to some of the participants.

Exploring computational ethics and explainable AI. According to one participant, we will have to design systems that are going to operate in environments where we cannot anticipate in advance all the things that could go wrong. Explainable AI and computational ethics are relevant for all places where AI systems are interacting with the physical world. As for computational ethics, AI researchers have begun establishing rules of their own. For example, some groups of technologists have created sets of ethical considerations.[6] In addition, researchers from six institutions recently formed a group called PERVADE (Pervasive Data Ethics for Computational Research), whose mission is to develop a clearer ethical process for big-data research for use by both universities and private companies. However, as one participant noted, the current and future developers of AI systems may operate by ethical standards or adhere to certain morals or values that may not be compatible with the rest of society or representative of those who will use the AI.

In conclusion, in our March 2018 technology assessment, we noted that AI technologies are already impacting a wide array of economic sectors. Our technology assessment also provides an overview of developments in the field of AI, focusing on the challenges, opportunities, and implications of these developments for policy making and research, and further helps clarify the prospects for the near-term future of AI and identifies areas where changes in policy and research may be needed.

Chairwoman Comstock, Chairman Weber and Ranking Members Lipinski and Veasey, this concludes my statement. I would be pleased to respond to any questions you or other Members may have.

GAO Contact and Staff Acknowledgments

If you or your staff have any questions about this testimony, please contact Timothy Persons at (202) 512-6522 or personst@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this statement. Individuals making key contributions to this testimony include Stephen Sanford (Assistant Director), Virginia Chanley (Analyst-in-Charge), and David Chrisinger. Key contributors to the prior work on which this testimony is based are listed in the product.

[6]Hila Mehr, "Artificial Intelligence for Citizen Services and Government," Ash Center for Democratic Governance and Innovation, Harvard Kennedy School, August 2017.

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

GAO's Mission	The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.
Obtaining Copies of GAO Reports and Testimony	The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (https://www.gao.gov). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to https://www.gao.gov and select "E-mail Updates."
Order by Phone	The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, https://www.gao.gov/ordering.htm. Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537. Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.
Connect with GAO	Connect with GAO on Facebook, Flickr, Twitter, and YouTube. Subscribe to our RSS Feeds or E-mail Updates. Listen to our Podcasts. Visit GAO on the web at https://www.gao.gov.
To Report Fraud, Waste, and Abuse in Federal Programs	Contact: Website: https://www.gao.gov/fraudnet/fraudnet.htm Automated answering system: (800) 424-5454 or (202) 512-7470
Congressional Relations	Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548
Public Affairs	Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800 U.S. Government Accountability Office, 441 G Street NW, Room 7149 Washington, DC 20548
Strategic Planning and External Liaison	James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707 U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548

Please Print on Recycled Paper.

Timothy M. Persons, Ph.D.
Chief Scientist
United States Government Accountability Office

Biography

Dr. Timothy M. Persons is the Chief Scientist of the United States Government Accountability Office as well as a Director for GAO's Center for Science, Technology, and Engineering (CSTE).

In these roles he is an expert on advanced analytics and cutting-edge science and technology, key highly-specialized national and international systems, engineering policies, best practices, and original research studies in the fields of engineering, computer, and the physical and biological sciences to ensure efficient, effective, and economical use of science and technology in government programs.

He has led and conducted advanced scientific and technical studies for the U.S. Congress on topics ranging from artificial intelligence, data analytics, bioforensics, oversight of high containment laboratories, 3D printing, nonmanufacturing, biodetection systems, homeland security imaging and nuclear detection systems, and freshwater conservation technologies.

Chairwoman COMSTOCK. Thank you. And I now recognize Mr. Brockman for five minutes.

TESTIMONY OF MR. GREG BROCKMAN, CO-FOUNDER AND CHIEF TECHNOLOGY OFFICER, OPENAI

Mr. BROCKMAN. Chairwoman Comstock, Chairman Weber, Ranking Member Lipinski, Ranking Member Veasey, members of both subcommittees, thank you for having me today to deliver testimony.

I'm Greg Brockman, Cofounder of OpenAI, a San Francisco-based nonprofit with a mission to ensure that artificial general intelligence, which we define as systems—the highly autonomous systems that outperform humans at most economically valuable work—benefits all of humanity.

Now, I'm here to tell you about the generality of modern AI, why AGI might actually be in reach sooner than commonly expected, and what action policymakers can take today.

So, first, what's OpenAI? We're a research company with one of the world's most advanced AI research and development teams. Yesterday, we announced major progress towards a milestone that we, Alphabet's subsidiary DeepMind, and Facebook have separately been trying to reach, which is solving complex strategy games which start to capture many aspects of the real world that were just not seen in board games like chess or Go.

We built a system called OpenAI Five, which learned to devise long-term plans and navigate scenarios far too complex to be programmed in by a human in order to solve a massively popular competitive game called Dota 2.

Now, in the past, AI-like technology was written by humans in order to solve one specific problem at a time. It was not capable of adapting to solve new problems. Today's AI, it's all based on one core technique, which is the artificial neural network, a single simple idea that, as it's run on faster computers, is proving to match a surprising amount of human capability. And this was in fact something that was shown in part by my fellow witnesses Dr. Li's work in image recognition. And artificial neural networks can be trained to perform speech recognition or computer vision. It just depends on the data that they're shown.

Now, further along the spectrum of generality is AGI. Rather than being developed for any one use case, AGI would be developed for a wide range of important tasks, and AGI would also be useful for noncommercial applications, including thinking through complex international disputes or city planning.

Now, people have been talking about AGI for decades, and so how should we think about the timeline? Well, all AI systems, they're built on three foundations. That's data, computational power, and algorithms. Next-generation AI systems are already starting to rely less on conventional data sets where a human has provided the right answer. For example, one of our recent neural networks learned by reading 7,000 books.

We also recently released a study showing that the amount of computation powering the largest AI training runs has been doubling every 3–1/2 months since 2012. That's a total increase of 300,000 times. And we expect this to continue for the next five

years using only today's proven hardware technologies and not assuming any breakthroughs like quantum or optical.

Now, to put that in perspective, that's like if your phone battery, which today lasts for a day, started to last for 800 years and then, five years later, started to last for 100 million years. It's this torrent of compute, this tsunami of compute. We've never seen anything like this. And so the open question is will this massive increase in combinational power, combined with near-term improvements in algorithmic understanding, be enough to develop AGI? We don't know the answer to this question today, but given the rapid progress that we are seeing, we can't confidently rule it out.

And so now what should we be thinking about today? What can policymakers be doing today? And so, you know, the first thing to recognize is the core danger of AGI is that it has fundamentally the potential to cause rapid change whether that's through machines pursuing goals that are mis-specified by their operator, whether it's through malicious humans subverting deployed systems, or whether it's an economy that grows in an out-of-control way for its own sake rather than in order to improve human lives.

Now, we spent two years. worth of policy research to create the OpenAI Charter, which in fact is a document I have right here in front of me. This contains three sections defining our views on safe and responsible AGI development. So that's—one is leaving time for safety and in particular refusing a race to the bottom on safety in order to reach AGI first. The second is to ensure that people at large rather than any one small group receive the benefits of this transformative technology. And the third is working together as a community in order to solve safety and policy challenges.

Now, our primary recommendation to policymakers is to start measuring progress in this field. We need to understand how fast the field is moving, what capabilities are likely to arrive when in order to successfully plan for AGI challenges. That moves towards forecasts rather than intuition. Measurement is also a place where international coordination is actually valuable, and this is important if we want to spread safety and ethical standards globally.

So thank you for your time, and I look forward to questions.

[The prepared statement of Mr. Brockman follows:]

Testimony of Mr. Greg Brockman

Chairwoman Barbara Comstock, Chairman Randy Weber, Ranking Member Daniel Lipinski, Ranking Member Marc Veasey, members of both subcommittees, thank you for having me today to deliver testimony on this important topic.

I'm Greg Brockman, co-founder of OpenAI, a non-profit artificial intelligence development organization. Our mission is to ensure that artificial general intelligence (AGI) — by which we mean highly autonomous systems that outperform humans at most economically valuable work — benefits all of humanity.

OpenAI

OpenAI has three main arms: capabilities, safety, and policy. All three of these areas, working in concert, are crucial to achieve our mission. Our capabilities arm is one of the most advanced AI research and development teams in the world. Our safety arm is responsible for developing techniques to ensure that AGI-level systems will operate as their human operators intend. Our policy arm is responsible for researching AGI's social challenges and providing information to policymakers.

On the capabilities side, one milestone in the field is solving complex strategy games[1], which capture many of the aspects of the real world not seen in previous milestones like Chess or Go. We recently announced OpenAI Five[2], a system we've created which has reached the semi-professional level at one of the most complex games played by humans, a ten-player team strategy video game called Dota. This system devises long-term plans and navigates scenarios far too complex to be programmed in by any human. We are aiming to play against top professionals during the Dota world championships in August. OpenAI Five taught itself the rules of the game by playing 180 years worth of games against itself each day. (For comparison, top human professionals have at least 12,000 hours of gameplay, so our system sees as many games each day as 100 human professionals have seen in their lifetimes.)

On the safety side, we recently developed a proof-of-concept technique[3] for allowing humans to monitor the behavior of advanced AI systems. We have also collaborated with Alphabet's subsidiary DeepMind to design AI systems which can learn from the implicit preferences of human trainers.

---

[1] Gershgorn, Dave. The massive global race to teach an AI to beat Starcraft II is under way. https://qz.com/1051052/deepmind-goog-and-facebook-fb-have-started-the-global-sprint-for-ai-to-beat-starcraft-ii/

[2] OpenAI Five. https://blog.openai.com/openai-five/

[3] Irving, Geoffrey, et al. AI Safety via Debate. https://blog.openai.com/debate/

On the policy side, we recently co-authored a report[4] forecasting how malicious actors could misuse AI, including recommendations of how to mitigate these threats. We're helping to develop the AI Index, an AI measurement and analysis initiative, as part of the Stanford One Hundred Year Study on AI. Our goal is to use this experience to make recommendations about how policymakers can measure and analyze the impact of AI on society. And we are attempting to nurture the field of AI policy to ensure there is a deep bench of thinkers about AI across all important actors — companies, research labs, and governments.

Narrow vs general AI

People often talk about narrow vs general AI in terms of whether they apply to one task (narrow) or many tasks (general). But there's also a dimension of competence: can they solve only easy tasks, or can they solve hard tasks? In practice, to build AI systems that solve harder problems, we've ended up creating increasingly general learning systems — since we let the machine learn more on its own rather than having a human provide knowledge or guidance.

Specifically, in the past, AI-like technology was written by humans to solve one specific problem. It wasn't capable of adapting to solve new problems.

In contrast, today's AI is all based on one core technique: the artificial neural network, in a form devised in the 1980s. This is a single, simple idea that is, as it scales, is proving itself to be able to match a surprising amount of human capability. Our neural networks still have a lot of room to grow — to give a sense of scale, though the numbers are not directly comparable, today they usually have around the same number of artificial neurons as an insect has biological ones.

In the 1980s, computers could only run tiny neural networks, so the resulting systems couldn't solve interesting problems. In 2012, computers were fast enough for a team of researchers (including my co-founder Ilya Sutskever) to train a large enough neural network to perform well on the task of categorizing images — performing far better than any other method. The neural network learned to categorize images by being shown many examples of already-categorized images, and this is now the dominant approach in the field rather than the previously hand-crafted rules (which were limited in performance). Since then, neural networks have become the standard tool for solving problems in a variety of fields such as speech recognition and machine translation.

To give you a sense of progress, here are some AI advances from recent years:

- **Image recognition:** AI's ability to correctly categorize images has gone from 75% (pre-neural network, 2011) accuracy to around 98% accuracy (neural network, 2017) on a difficult standard benchmark on which human accuracy is around 95%.

---

[4] Clark, Jack, et al. Preparing for Malicious Uses of AI. https://blog.openai.com/preparing-for-malicious-uses-of-ai/

- **Fake images & videos:** AI techniques are increasingly able to generate convincing fake images and videos — including fakes of politicians, such as Vladimir Putin and President Trump. In 2014, the best generated images were low-resolution images of fake people; by 2017, they were photorealistic faces that humans have trouble distinguishing from real ones[5]. Also in 2017, free software became available allowing people to create their own "deepfake" images.
- **Translation:** In 2014, researchers developed "neural machine translation" — where computers learn to translate between languages using only large datasets, lacking any of the specific rules which human translators use to do their work. In 2016, Google Translate performance significantly improved by switching to it; in 2017 Facebook improved its site translation by doing the same.
- **Speech recognition:** Due to switching to neural networks, over the past few years speech recognition went from barely working (such as we've all experienced when calling an automated phone tree) to running on smartphones with much higher accuracy.
- **Sophistication:** The complexity of games where neural network-based AIs can rival the top human players has increased in complexity from 1970s Atari games like Space Invaders or Breakout (2013) to rich strategic games like the board game Go (2015) to modern real-time strategy games like Dota in both 1-versus-1 (2017) and 5-versus-5 (2018) formats.

These advances, with neural networks at the core of all of them, are more general than past systems, and when trained properly, can achieve unprecedented performance at one or more interesting tasks. A system that can learn image recognition at record-setting levels could also learn to do the same in speech recognition. The tools used to generate fake images of politicians could also be used to synthesize new artistic paintings, or imaginary architectural plans. Neural machine translation systems can learn to translate between any pair of languages — provided we have the training data.

The next step along the spectrum are future AI systems that can accomplish very hard, valuable real-world tasks such as:

- automatically devising and performing scientific experiments in chemistry or neuroscience
- helping us design better or cheaper drugs, cars, computer software and hardware, and public infrastructure like transport or logistics systems
- performing surgery with more precision, safety, and efficiency than is possible for human surgeons, to
- orchestrating the movements of thousands of self-driving cars and drones across a city or rural area, whether to deliver ordinary goods or provide emergency supplies during an extreme weather event.

---

[5] Brundage, Miles, et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, page 15. https://arxiv.org/pdf/1802.07228.pdf

What is AGI?

AGI is even further along the spectrum of generality. Our working definition of AGI is systems that are sufficiently advanced that they can outperform humans at most economically valuable work — which includes starting companies, making business deals, and writing books. Such technology will need to be smart in a way unlike our traditionally literal, uncreative computers. The system's generality means it wouldn't be limited to commercial applications — it could also help with reasoning through complex international disputes, city planning, and even lawmaking or running countries. Rather than being developed for any one use-case, AGI would be developed for an entire spectrum of important tasks.

Why AGI could be developed sooner than commonly expected

AI systems are built on three foundations: algorithms, data, and compute (or amount of computational resource). Next generation AI systems being developed today are relying less on conventional datasets, since they can either consume freely-available unlabeled data (like a recent state-of-the-art language model we released which learns from an open dataset of books) or can expend compute to generate data. For example, by simulating a robot, we can create training data in quantities limited only by the number of computers available to run the simulation.

We recently released a study[6] showing that the amount of compute powering the largest AI training runs has been doubling every 3.5 months for the past six years (a total increase of 300,000x). This growth is significantly faster than the historic driver of hardware progress, Moore's Law, which had an 18 month doubling period (a 12x increase over the same period). AI compute progress is driven partly by faster computers, and partly by figuring out how to effectively train AI systems on many computers simultaneously. This means that data and compute are rapidly becoming less significant bottlenecks on AI progress.

We expect this trend to continue. We track over 45 hardware startups (most in the US) that are building next-generation AI computers. Most are building on proven technologies that do not require further breakthroughs like quantum or optical. As these computers hit the market, and as we figure out how to use many such computers at once, we expect the rate of breakthrough results to continue apace or even accelerate.

Our current computational paradigm allows for substantial increases in compute each year for at least the next five years. Will the incoming tsunami of compute (combined with near-term improvements in algorithmic understanding) be enough to develop AGI, or will we need to wait for some future algorithmic or hardware breakthrough? We don't know the answer to this question yet, but given today's rapid progress, it seems unwise to be too confident in either direction — at least before uncovering further evidence.

---

[6] Amodei, Dario, et al. AI and Compute. https://blog.openai.com/ai-and-compute

The post-AGI future

Investment in AI research is increasing rapidly due to how quickly AI advances can be deployed into products. Transformative applications on the horizon like self-driving cars promise to save lives, increase efficiency, and generate huge value, with the potential to add trillions of dollars to US GDP[7].

After AGI is created, we expect economic and technological growth to accelerate markedly — with the new growth driven primarily by teams of creative computers partnering with creative humans. We'll have the technological means to not just generate but also distribute essential resources (and hopefully much more) to ensure no one falls through the cracks, and will be able to concentrate on efforts like education, re-training, and re-skilling, to help people navigate the new economy. The benefits will be vast, and OpenAI believes those benefits should be equitably distributed, rather than locked up with any one entity.

Technological progress has been accelerating rapidly for the past few hundred years, and we expect the post-AGI world to add another jolt to the rate of progress. We should expect advances in curing disease, life extension, transportation and space travel, and communication.

Challenges of AGI development

Each stage of AI development will bring its own challenges.

Narrow AI challenges are easiest to understand and act on because they apply to existing systems. These are also the ones that today's corporations are most incentivized to fix. These challenges include issues such as fairness, transparency, privacy, and bias — all of which require serious attention if we want even more advanced AI technologies to have a positive impact. We expect narrow AI progress to increase the rate of technological progress across the board, further challenging today's policy machinery to keep pace; as we approach AGI, we expect the rate of change to increase further.

AGI challenges are harder to understand and foresee, partly because they apply to systems that have not yet been developed. OpenAI focuses on AGI challenges because we believe that they are simultaneously neglected and may happen sooner than is commonly believed.

The core danger with AGI is that it has the potential to cause rapid change. This means we could end up in an undesirable environment before we have a chance to realize where we're even heading. The exact way the post-AGI world will look is hard to predict — that world will likely be more different from today's world than today's is from the 1500s. Some open questions:

---

[7] Lanctot, Roger. Accelerating the Future: The Economic Impact of the Emerging Passenger Economy, page 5. https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/05/passenger-economy.pdf?cid=em-elq-26916&utm_source=elq&utm_medium=email&utm_campaign=26916&elq_cid=1494219

- What is the nature of international society in the post-AGI world? We've already seen technology amplify the impact that states and small groups have in the world. We expect AGI to further this trend.
- What will people do with their time as economic work becomes an increasingly smaller part of one's life? How can we help people live meaningful, enjoyable lives in such a world?
- How can AGI be deployed into our economic and social systems in a way that amplifies human preferences? Will all humans remain meaningful actors in society? Only humans in countries with powerful AGI deployments? Only the humans who own a share in the technology itself?
- How do we ensure that AGIs operate in line with the values of their operators? How do we avoid creating systems that cause social harm in blind pursuit of a poorly-specified goal — the technological version of what happened to Bear Stearns in 2008.

AGI will affect every sector of global society, and given the difficulty of these questions, we don't have the luxury of waiting to see how AGI starts affecting society before addressing its challenges. One example issue worth considering today is the possibility of a military arms race toward AGI. A military arms race would put pressure on deploying an AGI without adequately verifying that it is safe. AGI deployment will be challenging enough without pressure to gamble with safety. (Similar considerations also apply to pre-AGI AI technologies.)

Safe and responsible AGI development

Our views on safe and responsible AGI development are captured in three of the four sections of our Charter[8]: "Broadly Distributed Benefits", "Long-Term Safety", and "Cooperative Orientation".

- **Safety**. We do not yet know how hard it will be to make sure AGIs act according to the values of their operators. Some people believe it will be easy; some people believe it'll be unimaginably difficult; but no one knows for sure — which is why OpenAI believes that safety research is critically important. At the very least, any AGI project should leave enough time to get safety right. This includes taking steps — well in advance of the development of AGI – to avoid an uncoordinated race. In this vein, our Charter commits us to assisting rather than competing with a value-aligned, safety-conscious project that comes close to building AGI before we do.
- **Broadly Distributed Benefits**. AGI will create unprecedented economic benefits. If AGI can truly produce not just a Microsoft-sized amount of value, but 100 Microsofts or more, then returns beyond some point should not exclusively belong to a small group of people. The rest of humanity will have assumed the risks of developing and deploying AGI, and everyone deserves a fair share in the post-AGI future.

[8] OpenAI Charter. https://blog.openai.com/openai-charter/

- **Cooperative Orientation**. AGI has the potential to be the most socially beneficial technology humans ever create. The world is bigger than any one project, and any society which successfully builds safe AGI will win collectively. Thus, it's important that value-aligned AGI projects view themselves in "friendly competition". Today, we are all competing for talent and prestige. But we need the ability to come together under one roof in some form before building such a powerful system, bringing together companies (and hopefully governments) to ensure the resulting technology benefits everyone.

Policy recommendations

1. **Measurement**. Many other established voices in the field have tried to combat panic about AGI by instead saying it not something to worry about or is unfathomably far off. We recommend neither panic nor a lack of caution. Instead, we recommend investing more resources into understanding where the field is, how quickly progress is accelerating, and what roadblocks might lie ahead. We're exploring this problem via our own research and support of initiatives like the AI Index. But there's much work to be done, and we are available to work with governments around the world to support their own measurement and assessment initiatives — for instance, we participated in a GAO-led study on AI last year.
2. **Foundation for international coordination.** AGI's impact, like that of the Internet before it, won't track national boundaries. Successfully using AGI to make the world better for people, while simultaneously preventing rogue actors from abusing it, will require international coordination of some form. Policymakers today should invest in creating the foundations for successful international coordination in AI, and recognize that the more adversarial the climate in which AGI is created, the less likely we are to achieve a good outcome. We think the most practical place to start is actually with the measurement initiatives: each government working on measurement will create teams of people who have a strong motivation to talk to their international counterparts to harmonize measurement schemes and develop global standards.

It's easy to imagine the post-AGI world as a destination — but it is more of an arbitrary marker denoting a world with transformative AI technologies. There are many open questions around AGI, and the more we can understand where the field is, how fast we are moving, and what is likely to happen in upcoming years, the better prepared we will be to answer them. And perhaps the most important question AGI raises is that once the world has been fundamentally transformed by systems that perform tasks we'd historically thought of as "human" — what then?

Greg Brockman is co-founder and CTO of OpenAI, a non-profit artificial intelligence research company working to ensure that artificial general intelligence benefits all of humanity. He's an investor in over 30 startups and a board member of the Stellar digital currency system. He was previously the CTO of Stripe, a payments startup now valued at over $9B USD. Greg studied mathematics at Harvard and computer science at MIT.

Chairwoman COMSTOCK. Thank you. And we now recognize Dr. Li.

TESTIMONY OF DR. FEI-FEI LI,
CHAIRPERSON OF THE BOARD AND CO-FOUNDER, AI4ALL

Dr. LI. Thank you for the invitation, Congresswomen and Congressmen. My name is Fei-Fei Li. I'm here today as the Cofounder and Chairperson of AI4ALL, a national nonprofit organization focusing on bringing hands-on experience in AI research to high school students that have been traditionally underrepresented in the field of—in the STEM fields such as girls, people of color, and members of low-income communities. Our program began at Stanford University in 2015. This year, AI4ALL are expanded across North America to six university campuses.

I often like to share with my students that there's nothing artificial about artificial intelligence. It's inspired by people, it's created by people, and, most importantly, it has an impact on people. It's a powerful tool we're only just beginning to understand, and that's a profound responsibility.

I'm here today because the time has come to have an informed public conversation about that responsibility. With proper guidance, AI will make life better, but without it, it stands to widen the wealth divide even further, making technology even more exclusive, and reinforce biases we've spent generations trying to overcome. This will be an ethical, philosophical, and humanistic challenge, and it will require a diverse community of contributors. It's an approach I call human-centered AI. It's made of three pillars that I believe will help ensure AI plays a positive role in the world.

The first is that the next generation of AI technology must reflect more of the qualities that make us human such as a deeper understanding of the context we rely on to make sense of the world. Progress on this front will make AI much better at understanding our needs but will require a deeper relationship between AI and fields like neuroscience, cognitive science, and the behavior sciences.

The second is the emphasis on enhancing and augmenting human skills, not replacing them. Machines are unlikely to replace nurses and doctors, for example, but machine learning assistive diagnosis will help their job tremendously. Similar opportunities to intelligently augment human capabilities abound from health care to education, from manufacturing to agriculture.

Finally, AI must be guided by a concern for its impact. We must address challenges of machine biases, security, privacy, as well as at the society level. Now is the time to prepare for the effect of AI on laws, ethics, and even culture.

To put these ideas in practice, governments, academia, and industry will have to work together. This will require better understanding of AI in all three branches of government. AI is simply too important to be owned by private interests alone, and publicly funded research and education can provide a more transparent foundation for its development.

Next, academia has a unique opportunity to elevate our understanding and development of this technology. Universities are a

perfect environment for studying its effect on our world, as well as supporting cross-disciplinary next-generation AI research.

Finally, businesses must develop a better balance between their responsibility to shareholders and their obligations to their users. Commercial AI products have the potential to change the world rapidly, and the time has come to complement this ambition with ethical, socially conscious policies.

Human-centered AI means keeping humans at the heart of this technology's development. Unfortunately, lack of diverse representation remains a crisis in AI. Women hold a fraction of high-tech positions, even fewer at the executive level, and this is even worse for people of color. We have good reasons to worry about bias in our algorithms. A lack of diversity among the people developing these algorithms will be among its primary causes. One of my favorite quotes comes from technology ethicist Shannon Vallor, who says that "There's no independent machine values. Machine values are human values."

However autonomous our technology becomes, its impact on the world will always be our responsibility. With the human-centered approach, we can make sure it's an impact we'll be proud of. Thank you.

[The prepared statement of Dr. Li follows:]

I4ALL
AI4ALL
2148 Broadway
Oakland, CA 94612
contact@ai-4-all.org

House Science Committee Testimony

Dr. Fei-Fei Li

Introduction

As I often say to my students, there's nothing "artificial" about AI. It's inspired by people, it's created by people, and—most importantly—it impacts people. It's a powerful tool we're only just beginning to understand, and that's a profound responsibility. I'd like to talk about what this technology looks like today, the challenges we'll face in the coming years, and what we can do now to ensure we're building a future worth living in.

Artificial intelligence emerged in the mid 20th century as a quest to build machines with intellectual capabilities similar to those of the human mind. As a science, it draws on fields like cognitive science, neuroscience, statistics and mathematics. As a technology, it represents some of the most active developments in computer science and engineering, including machine learning (ML)—a family of techniques that use statistical modeling to learn from data. Today, AI and ML are part of a vibrant, interdisciplinary pursuit, with fields like robotics, natural language processing, computer vision, speech recognition, and even philosophy playing ever-growing roles.

It's also changing the world. Thanks to the recent convergence of three key factors—powerful algorithms, fast computing hardware, and the era of big data—AI has rapidly become a driving force in what some are calling the fourth industrial revolution. It can help drive our cars, assist with radiology diagnoses, optimize energy consumption, and even track deforestation from satellites. All told, AI spending is expected to rise from 2017's $12B to $57.6B by 2021, according to International Data Corporation (IDC).

We have good reason to be excited about AI. But this is a nascent field, and as a scientist, that humbles me. After all, we've never created a technology to mimic human qualities so closely, and we know little about the impact it will have on the world. Guiding its development will be an ethical, philosophical and humanistic challenge, and it will require a diverse community of contributors. I call this "Human-Centered AI", and it consists of three simple ideas.

Three Pillars of Human-Centered AI

Pillar I: AI must be more inspired by human intelligence

Much of AI's recent success lies in a narrow range of problems associated with an abundance of annotated example data—such as accurately labeled photos or text—analyzed by algorithms in process called *supervised learning*. But this approach has its limits. For instance, in my lab, an image-captioning algorithm once fairly summarized a photo as "a man riding a horse" but failed to note the fact that both were statues. Similar lapses are found in speech recognition, natural language processing, robotics and more.

Simply put, AI technologies such as machine perception remains shallow compared to our own visual intelligence. Algorithms lack our talent for extrapolating from few examples, they tend to miss the big picture when given disorganized information, and they're devoid of our remarkable capacity for context and emotion. Researchers are working hard to address these shortcomings, and progress will deliver tools that understand our needs with newfound depth. But their work underscores the growing need for collaboration between AI and neuroscience, cognitive science and the behavioral sciences.

Pillar II: AI should strive to enhance us, not replace us

A recent McKinsey report found that half of all current work activities can be theoretically automated by technology that *already* exists. This is an alarming fact, and there's no denying that job displacement will be a central challenge of the coming decade. But what if we imagine AI as a tool for *augmenting* human capabilities, rather than replacing them?

For example, consider the problem of hospital-acquired infections. Although they claim over 90,000 lives each year—many times more than traffic fatalities—their cause is often simple human error. In response, my Stanford colleagues and

collaborators are partnering with Stanford Children's Hospital to explore the use of smart sensors to track hand hygiene automatically, with the goal of improving health outcomes and giving caregivers greater peace of mind. So while a nurse or doctor's true skills are unlikely to be replaced by machines—complex problem solving in uncertain situations, bedside manner, and an emotional connection with patients—this research demonstrates how technology can enhance them.

Similar opportunities to intelligently augment human capabilities abound, from healthcare to education, manufacturing to search and rescue, and city planning to entertainment. It's true that automation is a challenge to certain forms of human labor, but it's also an opportunity to elevate our notion of work.

Pillar III: AI must be guided by a concern for its human impact

The sheer power of AI means we have an obligation to consider its unintended consequences as thoughtfully as its desired effects. This begins with the technology itself, where fairness is among the most pressing concerns. For example, a recent study revealed that widely used facial recognition software recognizes white, male faces with considerably higher accuracy than darker-skinned and/or female faces.

Addressing such inequities will force us to confront a host of related problems. One is the issue of *interpretability*, or an AI system's capacity to explain its reasoning. As they play a growing role in deliberations ranging from parole hearings to loan applications, this ability will be critical. And because AI relies on data to learn about the world through large quantities of images, text, video and personal records, privacy and security will pose major challenges as well.

The picture only gets more complex at the societal level. First, some degree of labor disruption is inevitable, and now is the time to prepare for its effects with proactive measures like retraining. Next, licensing and certification will have to expand to include technology that can perform medical, legal and other highly regulated tasks without human involvement. And our ethics, not to mention the laws that codify them, will have to evolve to answer questions of accountability as their role in our lives grows more consequential.

Society will soon face changes we can't even imagine, and these predictions are only the beginning. We can expect an exciting but uncertain future, and a sense of humility will serve us well.

What Can Be Done?

In Government

Like nuclear technology, biotech and energy security, AI should be an active topic of discussion among policymakers. This will require a far greater baseline of AI proficiency in all three branches of government, however, as well as ongoing dialogues with industry and academia. This technology is simply too important to be owned by private interests alone, and publicly-funded research can provide a transparent, ethical foundation for its development.

However, both public and private research will depend on a growing field of technologists with cutting-edge expertise in subjects like machine learning and data science. To ensure we can meet demand, it's vital we provide funding for education and diversity outreach in the STEM fields, with an emphasis on AI, starting as early as possible.

In Academia

Academia has a unique opportunity to bring clarity and rigor to our understanding of AI. The rare mix of intellectual resources found in our universities is a perfect environment for studying the effects of AI on our world, as well as supporting the cross-disciplinary demands of the next generation of research. Current campuses often separate computer science from other fields like the humanities and social sciences, but the time has come for this practice to change. Academic leaders can encourage discussions across departments in the form of classes, seminars, institutions, symposiums and joint projects.

It's encouraging to note that academic venues are already playing a unique role in combating bias. Organizations like FAT/ML (Fairness, Accountability, and Transparency in Machine Learning), IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, the Partnership on AI and AI Now Institute are active sources of thought leadership in the ethical implementation of AI. I hope to see this trend continue.

In Industry

Finally, businesses must develop a better balance between their responsibility to shareholders and their obligations to their users, community and society at large. After all, commercial products tend to change the world fastest. So while the tech industry's aggressive investment in AI promises rapid advances, the time has come to complement that ambition with ethical, responsible policies.

I'm happy to say some companies are already taking the first steps, with published guidelines intended to establish ethical boundaries and best practices for the field. But this is just a start. Across the tech industry and beyond, much more can be done to promote the responsible development of AI, including relationships with academia, nonprofits and governments to encourage the study of its human impact.

The Common Thread: More Humanity

Even after nearly two decades in this field, I'm still heartened by the creativity, intelligence and diversity of the people driving it. I introduce high school students to AI research through AI4ALL, a non-profit I co-founded, advise doctoral students at Stanford, and work with Google researchers worldwide. I've learned that no age, ethnicity or gender has a monopoly on innovation. In fact, more and more of the best work in our field comes from women, people of color, and other groups that have traditionally been underrepresented in STEM.

Unfortunately, a lack of representation remains a crisis in AI. A 2014 report revealed that women in tech feel stalled, encounter persistent bias, and are statistically more likely to leave their jobs. As a result, females hold only a third of all high-tech positions, and even fewer at the executive level. In fact, only 13% of AI companies have female CEOs. For people of color, the numbers are even worse.

The tech world has a long way to go to address its lack of diversity, starting with hiring practices. Improvements are likely to yield immediate benefit, as research[1] from McKinsey and Credit Suisse to the Journal of Personality and Social Psychology have quantified the benefits of diversity on workplace performance. But studies also suggests this problem

[1] https://hbr.org/2016/11/why-diverse-teams-are-smarter

has far earlier origins; girls, for instance, tend to lose interest in STEM curricula around age 15, usually due to a lack of peers and role models that look like them. How should we expect the world's most powerful algorithms to behave if they remain ignorant to the diversity of human experience?

Organizations like AI4ALL were founded to reverse this trend. Our summer program gives students from underrepresented communities first-hand experience in real AI labs at Stanford and UC Berkeley, and we're expanding across North America this year. 91% of our graduates feel confident they'll pursue a career in the field, and 97% feel they're part of a community in AI and computer science. In the words of one recent graduate, "I used to think I wasn't smart enough to do computer science and AI. But now I've gained so much confidence because of all the support, and being around other girls who are into the same thing."

Conclusion

As computational ethicist Shannon Vallor said, there are no independent "machine" values; machine values are *human* values. A human-centered approach to AI means these machines don't have to be our competitors, but partners in securing our wellbeing. However autonomous our technology becomes, its impact on the world will always be our responsibility. Now's the time to start making sure its an impact we can be proud of.

Dr. Fei-Fei Li Biography:

Fei-Fei Li is a Professor in the Computer Science Department at Stanford, and the Director of the Stanford Artificial Intelligence Lab. In 2017, she also joined Google Cloud as Chief Scientist of AI and Machine Learning. Dr. Li's main research areas are in machine learning, deep learning, computer vision and cognitive and computational neuroscience. She has published approximately 200 scientific articles in top-tier journals and conferences, including Nature, PNAS, Journal of Neuroscience, New England Journal of Medicine, CVPR, ICCV, NIPS, ECCV, IJCV, and IEEE-PAMI. Dr. Li obtained her B.A. degree in physics from Princeton in 1999 with High Honors, and her PhD degree in electrical engineering from California Institute of Technology (Caltech) in 2005. She joined Stanford in 2009 as an assistant professor, and was promoted to associate professor with tenure in 2012. Prior to that, she was on faculty at Princeton University (2007-2009) and University of Illinois Urbana-Champaign (2005-2006).

Dr. Li is the inventor of ImageNet and the ImageNet Challenge, a critical large-scale dataset and benchmarking effort that has contributed to the latest developments in computer vision and deep learning in AI. In addition to her technical contributions, she is a national leading voice for advocating diversity in STEM and AI. She is a co-founder of Stanford's renowned SAILORS outreach program for high school girls and the national non-profit AI4ALL. For her work in AI, Dr. Li has been a keynote or invited speaker at many conferences, including the World Economics Forum, the Grace Hopper Conference in 2017, and TED2015 main conference. She is a recipient of the 2017 Athena Academic Leadership Award, IAPR 2016 J.K. Aggarwal Prize, the 2016 NVIDIA Pioneer in AI Award, 2014 IBM Faculty Fellow Award, 2011 Alfred Sloan Faculty Award, 2012 Yahoo Labs FREP award, 2009 NSF CAREER award, the 2006 Microsoft Research New Faculty Fellowship, and a number of Google Research awards. Work from Dr. Li's lab have been featured in a variety of popular press magazines and newspapers including New York Times, Wall Street Journal, Fortune Magazine, Science, Wired Magazine, MIT Technology Review, Financial Times, and more. She was selected as a 2017 Women in Tech by the ELLE Magazine, a 2017 Awesome Women Award by Good Housekeeping, a Global Thinker of 2015 by Foreign Policy, and one of the "Great Immigrants: The Pride of America" in 2016 by the Carnegie Foundation (past winners include Albert Einstein, Yo-Yo Ma, Sergey Brin and more).

Chairwoman COMSTOCK. Thank you. And I now recognize myself for five minutes for questions.

Dr. Li, there's a generally accepted potential for AI-enabled teaching to a minimum to provide a backup for traditional classroom education. So as AI technology advances, it seems reasonable to assume that your traditional education, vocational training, homeschooling, and even college coursework will need to change and adapt. Could you maybe comment about how education in general and for specific groups and individuals might be transformed by AI and how we can make that positive and really sort of have more of a democratization of education, particularly higher education and in STEM and in science?

Dr. LI. Thank you for the question. Of course, I feel passionate about education. So I want to address this question in—from two dimensions. One is how could we improve the education of AI and STEM in general to more students and general community? Second is what can AI as a technology do to help education itself?

On the first dimension, as evidenced by our work in the AI4ALL, we really believe that it's simultaneously a crisis and an important opportunity that we involve more people in the development of AI technology. AI represents—humanity has never created a technology so similar or trying to resemble who we are, and we need AI to—we need technologists and leaders of tomorrow to represent this technology.

So, personally, I think we need to democratize AI's education to reach out to more students of color, girls, women from traditionally underrepresented minority groups. At AI4ALL for the past four years, we've already created an alumni population of more than 100 students, and through their own community and route outreach effort, we have touched the lives of more than 1,400 youth ranging from middle schoolers to high schoolers in disseminating this AI knowledge, and we need more of that in higher education.

The second dimension that I want to answer your question is AI as a technology itself can help improve education itself. In the machine learning community, I'm sure, Greg, you also agree with me that there is an increasing recognition of the opportunity for lifelong learning using technology as an assistive technology. I have colleagues at Stanford who focus on research and reinforcement learning and education of how to bring more technological assistance into the teaching and territorialization of education itself, and I think this could become a huge tool, as I was saying, to augment human teachers and human educators to—so that our knowledge can reach to more students and wider community.

Chairwoman COMSTOCK. Excellent. And for other witnesses, could you maybe comment on how academic institutions and industry could work with government on AI?

Mr. BROCKMAN. All right. So, you know, for—OpenAI's recommendation is really about starting with measurement, right, to really start to understand what's happening in the field. I think it's really about, for example, the study that we did showing the 300,000 times increase. We need more of that. We need to understand where things are going, where we are. I think the government is uniquely positioned to set some of the goalposts as well, and we've been pretty encouraged by seeing some of the work that

is happening at GAO and also DIUx has had some success with us. So we think it's really about starting a low touch way for the dialogue to start happening because I think right now the dialogue is not happening to the extent that it should.

Dr. PERSONS. All right. Thank you for the question. I do think that, as the committees have pointed out, this is a whole-of-society issue. It's going to be government in partnership with the private sector, with academia to look at things. So I think there is room for thought about how to learn by doing, creating internships and ways to try and solve real-world problems so that you have a mix of the classroom experience, as well as making and building—you'll fail a lot of course with these things but learning in a safe environment and then being able to grow expertise in that way.

Chairwoman COMSTOCK. Thank you. And, Dr. Li, did you have anything you wanted to add to that also? Okay. Well, thank you.

And I now will recognize Mr. Lipinski for five minutes.

Mr. LIPINSKI. Thank you. This is a fascinating topic. And there's—I want to try to move through some things quickly, but I'll get some good answers here.

It seems to me that, Mr. Brockman, you have a different view of how—of AGI, the possibilities of AGI and how quickly it can come, then, the GAO report. Is there a reason for this? Is there something you think that GAO is missing? And if Dr. Persons could respond to that.

Mr. BROCKMAN. So I don't know if I can comment directly on the report just not being familiar enough with all the details in there, but I can certainly comment on our perspective on AGI and its possibility. And a lot of it really comes down to, rather than—you know, I think that there's been a lot of more emotion or intuition-based argument. And to your opening remarks, you know, I think that science-based reasoning in order to project what's happening in this field is extremely important, and that something that we've spent quite a lot of effort on since starting OpenAI almost three years ago.

And so looking at the barriers to progress as compute data algorithms, data is something that's changing very rapidly in terms of what data we can use, the computation, the power there is growing at a rate that we've just never seen over the course of this decade. We're going to be talking, you know, I think about ten orders of magnitude, and that something where, if you were to compare that to the typical growth of compute, something like Moore's law, that the—over the period where we saw 300,000X increase in the past six years, we would've only seen 12X, right? That's a huge gap, and this is somewhere where we're sort of being projected into the future a lot faster than people realize.

Now, it doesn't mean that it's going to happen soon. It means that we can't rule it out. It means that for the next five years, as long as this hardware growth is happening, we're in a fog and it's hard to make confident projections. And so my position is that we can't rule it out. We know that this is—you know, we're talking about a technological revolution on the scale of the agricultural revolution, something that could be so beneficial to everyone in this world. And if we aren't careful in terms of thinking ahead and trying to be prepared, it really could be caught unaware.

Mr. LIPINSKI. And thank you. Dr. Persons, do you have a response on that?

Dr. PERSONS. Sure. I think—and with all respect for our Silicon Valley innovators who are upstarts and challenge the status quo, I think it's great that we have this system. The key thing that we're seeing is the convergence of these technologies that was mentioned by my panelists of the exponential power in computing, the ubiquitous nature of data, the sophistication of algorithms are all coming in.

But that said, many folks in the community are mildly skeptical about the rate at which general AI may come in this area because—for several reasons. One is just the way that we think about the problem now, the super complexity that is manifest in addressing the various challenges. You're looking at large data sets and looking at all the facets of them. It's much easier to say than to do.

And again, I think a lot of the—as you pointed out, the driving force here is the concern about general AI and taking over the world kind of thing, and it's just much harder to mimic human intelligence, especially in an environment where intelligence isn't even really defined or understood.

And I think, as Dr. Li pointed out, a lot of this is really about augmentation. That's something else we heard from our experts. It wasn't a replacement of humans; it was a how can we become better humans, more functional humans in doing these things? So a lot of it just gets down to the——

Mr. LIPINSKI. Let me—because I have a short time, sorry.

Dr. PERSONS. Thank you.

Mr. LIPINSKI. I just want to throw out quickly, the—there have been very different—vastly different opinions and—about the replacement of jobs and the disappearance of jobs and what the impact's going to be. Mr. Brockman, what do you think the impact will be?

Mr. BROCKMAN. So I think that with new technologies in the short term we always overestimate the degree to which they can make rapid change, but I think in the long-term that they do, I think technology is change in that we've seen with things like the internet, that there's been a lot of job displacement, both creation and destruction. And I think AI will be no different. I think the question of exactly which jobs and when I think we don't have enough information yet, and I think that that's where measurement really starts to come in. So we view it as an open question and a very important one.

Dr. PERSONS. And, sir, if I can just say as a bottom line, nobody really knows the impact on this, and of course our experts are saying to know more we might need to be able to encourage—let's—for example, our Bureau of Labor Statistics, a data-type agency that—out of the federal government to help provide more data or different data or things to help try and answer the question of what is the impact as this technology continues to unfurl.

That said, there's also a history of—when you—it goes back and attributed to Ned Ludd in the era of British industrialization and the concern of destroying the machines for the concern about loss of jobs, and yet—and many times throughout history, it's happened

in an array of technologies where net jobs actually increased. It just—they were more sophisticated jobs. They were toward higher value creation and more productivity. So there is hope with this technology as well.

Mr. LIPINSKI. And if the Chairwoman will allow, I want to hear from Dr. Li.

Dr. LI. I just want to say that technology inevitably throughout human civilization has an impact to change the landscape of jobs, but it's really, really critical, like my fellow panelists said, that we need to invest in the research of how to assess this change. It's not a simple picture of replacement, especially when this technology has a much greater potential and power to augment it.

I just spent days in the hospital ICU with my mother in the past couple of weeks, and—with my own health care and AI research, you recognize that a nurse in a single shift is doing hundreds of different tasks in our ICU unit where they're fighting for life and death for our patients, and these are not a simple question of replacing jobs but creating better technology to assist them and to make their jobs better and make the lives better for everyone. And that's what I hope we focus on, using this technology.

Mr. LIPINSKI. Thank you.

Chairwoman COMSTOCK. Thank you, Dr. Li. That's a wonderful example of really vividly explaining to us how that can be used because certainly, as we're an aging population in this country, that's a challenge we're all facing. And so the quality of life and improvement in each of those employees and nurses being able to do a better job, thank you for outlining that.

I now recognize Mr. Weber.

Mr. WEBER. Thank you, Madam Chair.

Dr. Li, is your mom okay? We hope that she is and pray and hope that she is okay.

Dr. LI. Thank you. I'm here. That means she's better.

Mr. WEBER. Okay. Otherwise, we were going to be missing two witnesses. Good.

Dr. LI. She's watching me right now.

Mr. WEBER. Well, good.

Chairwoman COMSTOCK. Hi, Mom. She's doing a great job.

Mr. WEBER. You're doing excellent. She's a proud mom, and that's some good medicine in and of itself right there.

Dr. LI. Thank you.

Mr. WEBER. So we're glad for that.

Dr. Brockman, you in your statement say that your mission was to actually make sure that artificial intelligence benefited people and was better for the most economically valuable work. Do you remember that?

Mr. BROCKMAN. So are——

Mr. WEBER. It's in your written statement.

Mr. BROCKMAN. That's right. So the—our definition of what AGI will be, whether created by us or anyone else but just the milestone is a system that can outperform humans and economically valuable work.

Mr. WEBER. Okay. Well, let me read it to you real quick. "I'm Greg Brockman, Cofounder, a nonprofit development organization. Our mission is to ensure that artificial general intelligence, by

which we mean highly autonomous systems that outperform humans at," quote, "most economically valuable work," end quote, "benefits all humanity." How would you define most economically valuable work?

Mr. BROCKMAN. So I think that, again—and, first of all, I just—you know, the question of—you know, AGI is something that the whole field has been working towards for—you know, really since the beginning of the field 50 years ago, and so the question of how to define it I think is something that is not entirely agreed-upon, that our definition is this—and when we think of it, we think of—you think about things like starting companies or very high intellectual work like that——

Mr. WEBER. Right.

Mr. BROCKMAN. —and, you know, also to things like going in cleaning up disaster sites or things that humans would be unable to do very well today.

Mr. WEBER. Okay. Well, I noticed that in your disagreement that Congressman Lipinski referred to with the report, and you call them Silicon Valley upstarts. At least you didn't call them young upstarts, so that's an advantage. Thank you for doing that. But you're literally looking at a new industry that, even though it's shifted—bless you—even though the shift is going to be changing, you're actually creating jobs for another industry.

And going back to Dr. Li's example with her mom in the IU talking about much the nurses do, how do you train for those jobs if it's moving as fast as you think it is?

Mr. BROCKMAN. Yes, and so, you know, one thing I think is also very important is that I don't think we have much ability to change the timeline to this technology. I think that there are a lot of stakeholders, there are a lot of different pieces of the ecosystem. And that—what we do is we step back and we look at the trends and we say what's going to be possible when. And so I think that the question of how to train—and again, that's going to be something—we're not the only ones that are going to have to help answer that question.

But I think that the place to start, it really comes back to measurement, right? If we don't know what's coming, if we can't project well, then we're going to be taken by surprise. And so, you know, I think that there are going to be lots of jobs and already have been created jobs that are surprising in terms of—you think about with autonomous vehicles, that we need to label all this data, we need to make sure that the systems are doing what we expect, and that all of that—that there's going to be humans that are going to help make these systems——

Mr. WEBER. But we would all agree, I hope—and this question is for all three panelists—all three witnesses, that the jobs they're going to create are well worth the transformation into all of that technology.

Dr. Persons, would you agree with that?

Dr. PERSONS. I would agree to that. I'll—let me give you a quick example if I may. Speaking with a former Secretary of Transportation recently, just a simple example of tollbooth collectors, we have now a system where you get the E–ZPass, you drive through, and you have less of a workforce there that could have had an im-

pact at that time for short period on the number or loss of jobs for tollbooth collectors, and yet it freed them up. It enabled them to perhaps do other things that were needed and large problems.

Mr. WEBER. Okay. And, Mr. Brockman, you were shaking your head. You would agree with that statement?

Mr. BROCKMAN. Absolutely. I think that the purpose of technology and improving——

Mr. WEBER. Sure.

Mr. BROCKMAN. —it is to improve people's lives.

Mr. WEBER. So, Dr. Li, I see you shaking your head, too?

Dr. LI. Yes, absolutely. In addition to the example Mr. Persons provided, I think deeply about the jobs that are currently dangerous and harmful for humans from fighting fires to search and rescue to, you know, natural disaster recovery. Not only we shouldn't put humans in harm's way if we can avoid it, but also we don't have enough help in these situations, and they are—this is where technology should be of tremendous help.

Mr. WEBER. Very quickly, I'm out of time, just yes or no. If we lose dominance in AI, that puts us in a really bad spot in worldwide competitiveness, would you agree?

Dr. PERSONS. Yes.

Mr. BROCKMAN. Yes.

Mr. WEBER. Yes. Thank you.

Dr. LI. Yes.

Mr. WEBER. Madam Chair, I yield back.

Chairwoman COMSTOCK. Thank you. Good question.

Now, I recognize Mr. Veasey for five minutes.

Mr. VEASEY. Thank you, Madam Chair.

We have heard about already from your testimony some of the advantages of AI and how it can help humankind, how it can help advance us as a nation and a country. But, as you know, there are people also that have concerns about AI. There's been a lot of sort of doomsday-like comparisons about AI and what the future of AI can actually mean.

To what extent do you think this scenario, this sort of, you know, worst-case scenario that a lot of people have pointed out about AI is actually something that we should be concerned about? And if there is a legitimate concern, what can we do to help establish a more ethical, you know, responsible way to develop AI? And this is for anybody on the panel to answer.

Mr. BROCKMAN. So I think thinking about artificial general intelligence today is a little bit like thinking about the internet in maybe the late '50s, right? If someone was to describe to you what the internet was going to be, how it would affect the world, and the fact that all these weird things were going to start happening, you're going to have this thing called Uber which you're going to be able to—you'd just—you'd be very confused. It'd be very hard to understand what that would look like and the fact that, oh, we forget to put security in there and that we'd be paying for that for, you know, 30 years' worth of trying to fix things. And now imagine that that whole story, which played out over really the course of the past 60, almost 70 years now was going to play out in a much more compressed timescale.

And so that's the perspective that I have when it comes to artificial general intelligence is the fact that it can cause this rapid change and it's already hard for us to cope with the changes that technology brings. And so the question of is it going to be malicious actors, is it going to be that the technology itself just wasn't built in a safe way, or is it just that the deployment that who owns it and the values that it's given aren't something that we're all very happy with? All of those I think are real risks, and again, that's something that we want to start thinking about today.

Dr. PERSONS. Thank you, sir. So I agree with that. I think the key thing is being clear-eyed about what the risks actually are and not necessarily being driven by the entertaining and yet this science-fiction-type narrative sometimes on these things projecting or going to extremes and assuming far more than where we actually are in the technology.

So it's—there are risks. It's understanding the risks as they are, and there are always contextual risks. Risks in automated vehicles are going to be different than risks in this technology in financial services, let's say. So it's really working, again, symbiotically with the community of practice and identify what are the things there? What are the opportunities? And there's going to be opportunities. And then what undesirable things do we want to focus on and then optimize from there on on how to deal with them. Thank you.

Mr. VEASEY. Mr. Brockman, in your testimony you had referenced a report outlining some malicious actors in this area. Could you sort of elaborate on some of your findings in these areas?

Mr. BROCKMAN. That's right. So OpenAI was a collaborator on this research report projecting not necessarily today what people are doing but looking forward what are some of the malicious activities that people could use AI for. And so that report—let's see. Yes, I think maybe the most important things here you start thinking about a lot of things around information, privacy, the question of how we actually ensure that these systems do what the operator intends, despite potential hacking. You think about autonomous systems that are taking action on behalf of humans that are subverted and whether, again, it's—you know, that this report focuses on active action. You think about autonomous vehicles and if a human hacker can go and take control of a fleet of those, some of the bad things that could happen.

And so, you know, I think that this report should really be viewed as we need to be thinking about these things today before these are a problem because a lot of these systems are going to be deployed in a large-scale way, and if you're able to subvert them, then, you know, the—all of the problems that we've seen to date are going to start having a very different flavor where it's not just privacy anymore; it's also systems that are deployed in the real world that are actually able to affect our own well-being.

Mr. VEASEY. Thank you. Madam Chair, I yield back.

Chairwoman COMSTOCK. Thank you. And I now recognize Mr. Rohrabacher.

Mr. ROHRABACHER. Thank you very much, Madam Chairman.

This, as in all advances in technology, can be seen as the great hope for making things better, or the new idea that there might be new dangers involved, or the new technologies will help certain

peoples but be very damaging to others. And I think that where that fear would be most recognizable is in terms of employment and how in a free society people earn a living. And are we talking about here about the development of technology that will help get the tedious and remedial or the lower-skilled jobs that can be done by machine, or are we talking about the loss of employment by machines that are designed to really perform better than human beings perform in high-level jobs? What are we talking about here?

Dr. LI. Okay. So I can—I'm still going to use health care as an example because I'm familiar with that area of research. So if you look at recent studies by McKinsey and other institutions on employment and AI, there is a recognition that we need to talk a little more nuanced than just entire job but the tasks under each job. The technology has the potential to change the nature of different tasks. Again, for example, take nurse—a job of a nurse as an example. It—no matter how rapidly we develop the technology and the most optimistic assessment, it's very hard to imagine the entire profession of nurse—nursing would be replaced, yet within the nursing jobs there are many opportunities that certain tasks can be assisted by AI technology.

For example, a simple one that costs a lot of time and effort in nursing jobs is charting. Our nurses in our, again, ICU rooms, our patient rooms spend a lot of time typing and charting into a system, into a computer while that's time away from patients and other more critical care. So these are the kind of tasks under a bigger job description that we can hope to use technology to assist them and augment——

Mr. ROHRABACHER. So are we talking about robots here or a box that thinks and is able to make decisions for us? What are we talking about?

Dr. LI. So AI technology is a technology of many different aspects. It's not just robot. In this particular case, for example, natural language, understanding the speech recognition, and possibly in the form of a voice assistant would help charting. But maybe delivering of simple tools on the factory floor will be in the form of a small simple delivery robot. So there are different forms of machines.

Mr. ROHRABACHER. I see. Well, there are many dangerous jobs that I could see that we'd prefer not having human life put at risk in order to accomplish the goal. And, for example, at nuclear power plants it would be a wondrous thing to have a robotic response to something that could cause great damage to the overall community but would kill somebody if they actually went in to try to solve a problem. And I understand that and also possibly with communicable diseases where people need to be treated but you're putting people at great risk for doing that.

However, with that said, when people are seeking profit in a free and open society, I would hate to think that we're putting out of work people with lower skills, and we need the dignity of work and of earning your own way once we know now that when you take that away, it really has a major negative impact on people's lives.

So I want to thank you all for giving us a better understanding of what we're facing on this, and let's hope that we can develop this technology in a way that helps the widest variety of people and not

just perhaps a small group that will keep their jobs and keep the money. So thank you very much.

Chairwoman COMSTOCK. Thank you. And I now recognize Ms. Bonamici for five minutes.

Ms. BONAMICI. Thank you so much. Thank you to our witnesses.

First, I want to note that our nation has some of the best scientists and researchers and engineers in the world, but without stronger investments in research and development, especially long-term foundational research, we risk falling behind, especially in this important area. I hope the research continues to acknowledge the socioeconomic aspects as well of integrating AI technologies.

In my home State at the University of Oregon we have the Urbanism Next center. They're doing some great work bringing together interdisciplinary perspectives, including planning and architecture and engineering and urban design and public administration with public, private, and academic sectors to discuss how leveraging technology will shape the future of our communities. Their research has been talking about emerging technologies like autonomous vehicles and the implications for equity, health, the economy, and the environment and governance.

Dr. Persons, can you discuss the value of establishing this type of partnership between industry, academia, and the private sector to help especially identify and address some of the consequences intended and unintended of AI as it becomes more prevalent? And I do have a couple more questions.

Dr. PERSONS. Sure, I'll answer quickly. The short answer is yes. Our experts and what we're seeing is the value in public-private partnerships because, again, it would be a mistake to look at this technology in sort of isolated stovepipes, and it would need to be an integrated approach to things, the federal government having its various roles but key—like your mentioning at University of Oregon, key academic and research questions. There's many, many things to research and questions to answer and then of course industry, which has an incredible amount of innovation and thinking and power to drive things forward.

Ms. BONAMICI. Terrific. Thank you. Dr. Li, I have a couple questions. You discuss the labor disruption, and I know that's brought up a couple of times and the need for retraining. We really have sort of a dual skills gap issue here because we want to make sure there are enough people who have the education needed for the AI industries, but we also are talking about the workers like you mentioned, the workers in tollbooths who will be displaced. But with the rapid development of technologies and the changes in this field, what knowledge and skills are the most important for a workforce capable of addressing the opportunities and the barriers to the development?

I serve on the Education and Workforce Committee, and this is a really important issue is how do we educate people to be prepared for such rapid changes?

Dr. LI. So AI is fundamentally a scientific and engineering discipline, and to—as an educator, I really believe in more investment in STEM education from early age on. We look at—in our experience at AI4ALL when we invited these high school students in the age of 14, 15, 16 to participate in AI research, their capabilities

and potential just amazes me. We have high school students who have worked in my lab and won best-paper award at this country's best AI academic conferences. And so I believe passionately that STEM education is critical for the future for preparing AI workforce.

Ms. BONAMICI. Thank you. And as everyone on this committee knows, I always talk about STEAM because I'm a big believer in educating both halves of the brain, and students who have arts education tend to be more creative and innovative.

Also, Dr. Li, in your testimony you talk about how AI engineers need to work with neuroscientists and cognitive scientists to help AI systems develop a more human feel. Now, I know Dr. Carbonell is not here today, but I noted in his testimony he wrote, "AI is the ability to create machines who perform tasks normally associated with human intelligence." I'm sure that was an intentional choice to humanize the machine, but I wanted to ask you, Dr. Li, about—he's not here to explain, but I have no doubt that was intentional. In your testimony you talk about the laws that codify ethics. How is this going to be done? Can you go into more depth about how would these laws be done? Who would determine what is ethical? And would it be a combination of industry, government determining standards? How is—how are we going to set the stage for an ethical development of AI?

Dr. LI. Yes, so thank you for the question. I think for technology as impactful as AI is to human society, it's critical that we have ethical guidelines. And different institutions from government to academia to industry will have to participate in this dialogue together and also by themselves.

Ms. BONAMICI. Are they already doing that, though? You said they'll have to but is somebody convening all of this to make sure that there are——

Dr. LI. So there are efforts. I'm sure Greg can add to this. Industries in Silicon Valley we're seeing company starting to roll out AI ethical principles and responsible AI practices in academia. We see that ethicists and social scientists coming together with technologists holding seminars, symposiums, classes to discuss the ethical impact of AI. And hopefully, government will participate in this and support and invest and these kind of efforts.

Ms. BONAMICI. Thank you. I see my time is expired. Thank you, Madam Chair. I yield back. Oh, Mr. Chairman, thank you.

Mr. WEBER. [Presiding] I thank the gentlelady.

And the gentlelady from Arizona is recognized for five minutes.

Mrs. LESKO. Thank you, Mr. Chair.

I want to thank the testifiers today, very interesting subject and something that kind of spurs the imagination about science fiction shows and those type of things.

What countries are the major players in AI, and where does the United States rank in competition with them? And that's to any panelist or all panelists.

Mr. BROCKMAN. So, you know, today, I think that the United States actually ranks possibly top of the list. You know, I think there are lots of other countries that are investing very heavily. You know, China is investing heavily, lots of countries in Europe are investing heavily. The—you know, DeepMind is subsidiary of

a U.S. company but located in London. And I think that, you know, it's very clear that AI is going to be something of global impact, and I just think the more that we can understand what's happening everywhere and figure out how we can coordinate on safety and ethics in particular, the better it's going to go.

Dr. PERSONS. Yes, I—thank you for the question. I think wherever there is large amounts of computing, large amounts of data, and a strong desire to innovate and continue to develop again in this sort of fourth industrial revolution that we're moving on, then you—it drives toward certainly China and then our allies and colleagues in Western Europe and developed worlds. Thank you.

Mrs. LESKO. And is there—did you want to answer?

Mr. BROCKMAN. Sorry——

Mrs. LESKO. Go ahead.

Mr. BROCKMAN. If I could just add that, you know, the most important thing to continue to lead in the field, it's really about the talent. And right now, we're doing a great job of bringing all the talent in. At OpenAI we have a very wide mix of national backgrounds and origins, and I think as long as we can keep that up, that we'll be in very good shape.

Mrs. LESKO. Thank you. And, Mr. Chair, I have one more question, and I think this has been asked in different ways before, but what steps are we guarding against, espionage from, let's say—you said China is involved in this and that's basically my question—espionage, hacking, those type of things. What guidelines are currently taking place, and who's preventing this? Is it the private companies themselves? Is government involved? Thank you.

Mr. BROCKMAN. So one thing that's a very atypical about this field is because it really grew out of an academic—very small number of academic labs that the overarching ethos in the field is actually to publish. And so all of the core research and development is actually being shared pretty widely. And so I think that as we're starting to build these more powerful systems and this is one of the parts of our charter that we need to start thinking about safety and keeping—you know, thinking about things that should not be shared, and so I think that this is a new muscle that's being built. It's right now kind of up to each company, and I think that that something that we're all starting to develop. But I think having a dialogue around what's okay to share and what things are kind of too powerful and should be kept private, that's just the dialogue that's starting now.

Dr. PERSONS. And certainly IP or intellectual property protection is a critical issue. I think of one former Director of the National Security Agency mentioned that we're—at the time it was unprecedented theft of U.S. intellectual property at that time just because of the—it's the blessing and curse of the internet. It's a blessing it's this open and the curse is it's open. And so AI is going to I think be in that category.

In terms of what's being done in terms of cybersecurity, it is something our experts pointed out and said it is an issue. As this Committee well knows, it's easier said than done, and who has jurisdictions in the U.S. federalist system about particularly a private company and protection of that, the role of the federal government versus the company itself in an era where, as I think Mr.

Brockman has pointed out, is sort of the big-data era where data are the new oil, yet we want to be open at the same time so that we can innovate. So managing that dialectical tension is going to be a critical issue, and there's no easy answer.

Mrs. LESKO. Thank you. Mr. Chair, I yield back.

Mr. WEBER. The Chair recognizes Ms. Esty for five minutes.

Ms. ESTY. Thank you, Mr. Chair, and I want to thank the witnesses for this extremely informative and important conversation that we're having here today.

I hail from the State of Connecticut where we see a lot of innovation at UCONN, at Yale, at lots of spinoffs on the sort of narrow AI question. But I think for us really the issue is more about that general AI. And, Mr. Brockman, your discussion of the advances, which makes Moore's law look puny in comparison, is really where I want to take this conversation about, Dr. Li, your discussion, which I think is incredibly important, about diversity. We saw what happened to Lehman Brothers by not being diverse. I am extremely concerned about what the implications are for teaching a—as it were, if it's garbage in, it's going to be garbage out. If it's a very narrow set of parameters and thought patterns and life experiences that go into AI, we will get very narrow results out. So, first, I want to just talk—get your thoughts on that.

And the second is on this broader ethical question. We've looked for many years—I remember back when I was a young lawyer working on bioethical issues. The Hastings Center got created to begin to look at these issues. This Committee has been grappling with CRISPR and the implications with CRISPR. I think about this being very similar, that AI has many similar implications for ethical input.

So if you can opine on both of those questions and recognize we have got two minutes—three minutes left—about both the ethical—whether we need centers to really bring in ethicists, as well as technologists, and then the importance of diversity on the technology side so that we get the full range of human experience represented as we're exciting—our exciting new entry into this fourth revolution. Thanks.

Dr. LI. Yes, in fact when—just now—thank you for asking that question. Just now when somebody is using the term doomsday scenario, to me I think if we wake up 20 years from now, whatever years, and we see the lack of diversity in our technology and leaders and practitioners, that would be my doomsday scenario. So it's so important and critical to have diversity for the following three reasons, like you mentioned. One is shared jobs that we're talking about. This is a technology that could have potential to create jobs and improve quality of life, and we need all talents to participate in that.

Second is innovation and creativity like you mentioned in Connecticut and other places. We need that kind of broad talent to add in the force of AI development.

And the third is really justice and moral values, that if we do not have this wide representation of humanity representing the development of this technology, we could have face-recognition algorithms that are more accurate in recognizing a male—white male faces. And these are—we could have dangers of out—biased algo-

rithms making unfair loan application decisions. You know, there are many potential pitfalls of a technology that's biased and not diverse enough.

Which brings us to this conversation and dialogue of ethics and ethical AI. You're right. Previous disciplines like nuclear physics, like biology have shown us the importance of this. I don't know if there is a single recipe, but I think the need for centers, institutions, boards, and government committees are all potential ways to create and open this dialogue. And again, we're starting to see that, but I think you're totally right. It's critical issues.

Ms. ESTY. Mr. Brockman?

Mr. BROCKMAN. If I may, so I agree completely with my fellow witness. So diversity is crucial to success here. So actually—so we have a program called OpenAI scholars where we brought in a number of people from underrepresented backgrounds into the field and provided mentorship and they're working on projects and spinning up. One thing that we found that I think is very encouraging is it's actually very easy to take people who do not have any AI or machine learning background and to make them into extremely productive first-class researchers and engineers very quickly. And that's, you know, one benefit of this technology being so new and nascent is that in some ways it—we're all discovering as we go along, too, so becoming an expert, there just isn't that high of a bar. But—so I think that everyone putting effort in to places where the expertise is, I think it's on them to make sure that they're also bringing in the rest of the world.

On the ethical front, that's really core to my organization. That's the reason we exist, that we do think that, you know, for example, when it comes to the benefits of who owns this technology, who gets it—you know, where did the dollars go, we think it belongs to everyone. And so one of the reasons I'm here is because I think that this shouldn't be a decision that's made just in Silicon Valley. I don't think that the question of the ethics and how this is going to work should be in the hands solely of people like me. I think that it's really important to have a dialogue, and again, that something where, you know, I hope that that will be one of the outcomes of this hearing.

Ms. ESTY. Thank you very much.

Mr. WEBER. The gentleman now recognizes Mr. McNerney.

Mr. MCNERNEY. Well, I thank the Chair for holding this and the Ranking Member, and I thank the witnesses, really very interesting testimony and diverse in its own right.

One of the things that I think that's important here is—with this committee is how does the government react to AI? Do we need to create a specific agency? Does that agency report to Congress or to the Administration? Those sorts of things I think are very important.

Dr. Brockman, you said—I think one of the most important things was that we need a measure of AI progress. Do you have a model or some description of what that would look like?

Mr. BROCKMAN. Yes, I do. Thank you for the question. And so, first of all, I don't think that we need to create new agencies for this. I think that existing agencies are well set up for this. And I

was actually very encouraged to hear that people are talking about giving NIST a remit to think about these problems.

Again, GAO and DIUx are already starting to work on this. For example, DIUx had a satellite imagery data set, hosted a public competition. The kind of thing that we think would be great for government to do as well is to have standardized environments where academics and private sector can test robotic approaches, setting up competitions towards specific problems that various agencies and departments want to be solved. All of those I think can be done without any new agency, and I think that that's something that you can both get benefits directly to the relevant agencies, also understand the field, and also start to build ties between private sector and public sector.

Mr. MCNERNEY. I'm one of the founders of the Grid Innovation Caucus. What are the most likely areas we'll see positive benefits to the grid, to electric grid stability and resiliency? Who would be the best to answer that? Mr. Persons?

Dr. PERSONS. Sure. Thank you for the question. I think one of the ways—GAO has done a good deal of work on this issue, but it's just protection of the electrical grid in the cybersecurity dimension. So as one of our scenarios or profiles that we did in this report, what our experts and what folks are saying, and again based on the leadership of this Committee and the importance of cyber is that it's a—without which nothing—AI is going to be a part of cyber moving forward, and so protection of the grid in the cyber dimension is there.

Also, I think, as the Chairman mentioned earlier, the word optimization, so how we optimize things and how algorithms might be able to compute and find optimums faster and better than humans is an opportunity for grid management and production. Thank you.

Mr. MCNERNEY. So AI is also going to be used as a cyber weapon against infrastructures or potentially used as a weapon, is that right?

Dr. PERSONS. There are concerns now when you look at a broad definition of AI and you look at bots now that are attacking networks and doing distributed—what are—DDOS or distributed denial of service attacks and things like that, that exists now. You could—unfortunately, in the black hat assumption you're going to assume that as AI becomes more sophisticated, the white hats, and so, too, unfortunately, the black hat side of things, the bad guys are going to also become more sophisticated. And so it's going—that's going to be the cat-and-mouse game I think moving forward.

Mr. MCNERNEY. Another question for you, Dr. Persons. In your testimony you mentioned that there's considerable uncertainty in the jobs impact of AI.

Dr. PERSONS. Yes.

Mr. MCNERNEY. What would you do to improve that situation?

Dr. PERSONS. Our experts were encouraging specific data collected on this. Again, we have important federal agencies like BLS, Bureau of Labor Statistics, that work on these issues, what's going on in the labor market, for example, and it may just be an update to what we collect, what questions we ask as a government, how we provide that data, which of course is very important to our understanding of unemployment metrics and so on.

So there are economists that have thoughts about this. We had some input on that. There's no easy answer at this time, but the idea that there is an existing agency doing that sort of thing is there. The key question is how could we ask more or better questions on this particular issue on artificial systems?

Mr. MCNERNEY. Thank you. Dr. Li, you gave three conditions for progress in AI being positive. Do you see any acceptance or general wide acceptance of those conditions? How can we spread the word of that so that the industry is aware of them and the government is aware of them and that they follow those sorts of guidelines?

Dr. LI. Thank you for asking. Yes, I would love to spread the word. So I think I do see that—the emergence of efforts in all three conditions. The first one is about more interdisciplinary approach to AI and ranging from universities to industry, we see the recognition of neuroscience, cognitive science to cross pollinate with AI research.

I want to add we're all very excited by this technology, but as a scientist, I'm very humbled by how nascent the science is. It's only a science of 60 years old compared to traditionally classic science that's making human lives better every day, physics, chemistry, biology. There's a long, long way to go for AI to realize its full potential to help people. So that recognition really is important, and we need to get more research and cross-disciplinary research into that.

Secondly is the augmenting human, and again, a lot of academic research as well as industry startup efforts are looking at assistive technology from disability to, you know, helping humans. And the third is what many of us focus on today is the social impact from studying it to having a dialogue to having—to working together through different industry and government agencies. So all three are the elements of human-centered AI approach, and I see that happening more and more.

Mr. MCNERNEY. Thank you. I yield back.

Mr. WEBER. The Chair now recognizes the gentleman from New York. No. The Chair now recognizes the gentleman that's not from New York, Mr. Palmer.

Mr. PALMER. Thank you, Mr. Chairman. I'd like to know if AI can help people who are geography-challenged.

Mr. WEBER. The gentleman's time has expired.

Mr. PALMER. I request that that question and response be removed from the record.

I do have some questions. In my district, we have the National Computer Forensics Institute, which deals with cybercrime, and what I'm wondering about is with the emergence of and evolution of AI. What are we putting in place because of the potential that that creates for committing crime and for solving crime? Dr. Persons, do you have any thoughts on that?

Dr. PERSONS. Well, certainly in one of the areas we did—thank you for the question. One of the areas we did look at in general was just criminal justice. So, I mean, just the risks that are there in terms of the social risks, making sure the scales are balanced exactly as they ought to be, that justice is blind, and so on was the focus of that.

However, I think in terms of the criminal forensics, AI could be a tool that helps suss out what—you know, in a retrospective sense what happened. But again, it's an augmentation that's helping the forensic analyst who would know what things look like. And the algorithm would need—in the machine-learning sense of things would need to learn what the risks might be going forward so that you perhaps could identify things more proactively and perhaps in near or at real-time. So that's the opportunity for this. Again, AI is a tool and cyber was a key message we heard moving forward.

Mr. PALMER. Any thoughts on that?

Mr. BROCKMAN. So, today, you know, we're already starting to see some of the security problems with the methods that we're creating, for example, that there's a new class of attack called adversarial examples where researchers are able to craft like a physical patch that you could print out and put on any object. They'll make a computer vision system think that it's whatever object you want it to be, so you could put that on a stop sign and confuse a self-driving car, for example. So these sorts of ways of subverting these powerful systems is something we're going to have to solve and going to have to work on, just like we've been working on computer security for more conventional systems.

And I think that the way to think about if you could successfully build and deploy an AGI, what that would look like, in many ways it's kind of like the internet in terms of being very deeply integrated in people's lives but also having this increasing amount of autonomy and representation and taking action on people's behalf. And so you'll have kind of this question of how do you make sure, you know, first of all, that's something that could be great for security if these systems are well-built and have safety in their core and are very hard to subvert. But also if it's possible for people to hack them or to cause them to do things that are not aligned with the value of the operator, then I think that you can start having very large-scale disruption.

Mr. PALMER. It also concerns me in the context of—it was announced a couple of weeks ago that the United States plans to form a space corps. We know that China has been very aggressive in militarizing space. If you have any thoughts on that discussion of how artificial intelligence will be used in regard to space. Communication systems that are highly vulnerable already, I think that there's some additional vulnerability that would be created. Any thoughts on that? And any one of the three of the panelists.

Dr. PERSONS. Yes, sir. So in terms of the risk in space, obviously, one of the key concerns for AI is weaponization and—which I think is part of that and so much less the space domain or any other one. And so I know our Defense Department has key leadership thinking on this and working strategically on how do we operate in an environment where we have to assume there's going to be the adversary that might not operate in the ethical framework that we do and to defeat that, but there's no simple answer at this time other than our Defense Department is thinking about it and working on it.

Mr. PALMER. And he's not here obviously to testify, but Dr. Carbonell's testimony, he made a statement that we need to produce or AI researchers, especially more U.S. citizen or perma-

nent resident AI researchers. And I think that kind of plays into that issue of how do we deal with AI in space. That's one of the reasons why I've been pushing for a college program like an ROTC program to recruit people into the space corps in these areas, start identifying students when they're maybe even in junior high and scholarship them through college to get them into these positions. Any thoughts on that?

Dr. PERSONS. I'll just answer quickly and just say I think, as Dr. Li has I think elegantly pointed out before, this is really an interdisciplinary thing. I think there's going to be a need for sort of the STEM, STEAM specialist that's particularly focused on this, but I think any particular vocation is going to be impacted in one way or the other, just like you could imagine rewinding a couple decades or a few decades—I'll date myself—but when the advent of the personal computer, the PC coming in and how that affected. Now, we walk into any vocation and somebody's using a PC or something like that and it's not unusual, but at the time you had to learn how to augment yourself or your task with that. And I think that will be the case.

Mr. PALMER. Well, we're—if I may, Mr. Chairman—just to add this final thought that we've had to deal with some major hacks, federal government systems that are hacked, and what we're faced with, we're competing with the private sector for the best and brightest in terms of cybersecurity. We're going to find ourselves in the same situation with AI experts, the truly skilled people, and that's why I'm suggesting that we may need to start thinking about how do we recruit these people and get them as employees of the federal government? And that was my thoughts on setting up an ROTC-type program where we would recruit people in, we'd scholarship them, whether it's for cybersecurity or for AI and with a four- or five-year commitment to work for the federal government because there's going to be tremendous competition. And the federal government has a very difficult time competing for those type people.

So with that, Mr. Chairman, I yield back.

Mr. WEBER. Now, the Chair recognizes the gentleman from New York.

Mr. TONKO. It's okay. We're patient. I thank our respective Chairs and Ranking Members for today's very informative hearing.

And welcome and thanks to our witnesses.

I'm proud to represent New York's 20th Congressional District where our universities are leading the way in artificial intelligence research and education initiatives. SUNY Polytechnic Institute is currently the home of groundbreaking research developing neuromorphic circuits which could be used for deep learning such as pattern recognition but are also useful for AI or machine learning.

In addition, the institute has established an ongoing research program on restive memory devices. Rensselaer Polytechnic Institute, RPI, is pushing the boundaries of artificial intelligence in a few different areas. In the healthcare front, RPI is focusing on improving people's lives and patient outcomes by collaborating with Albany Medical Center to improve the performance of their emergency department by using AI and analytics to reduce the recur-

rence of costly ER visits by patients. And RPI researchers are also collaborating with IBM to use the Watson computing platform to help people with prediabetes avoid developing the disease.

In our fight to combat climate change and protect our environment, researchers at RPI and Earth and Environmental Science are working with computer science and machine learning researchers to apply cutting-edge AI to climate issues. In the education space, RPI is exploring new ways to use AI to improve teaching, as well as new approaches to teaching AI and data science to every student at RPI.

With all that being said, there are tremendous universities across our country that are excelling in AI research and education. And what are some of the keys to helping AI institutions like them to excel? What do we need to do? What would be the most important? That's to any one of our panelists.

Dr. LI. So thank you for asking this question. I think, just like we recognize AI really is such a widespread technology that I think one thing to recognize is that it is still so critical to support basic science research and education in our universities. This technology is far from being done. Of course, the industry is making tremendous investment and effort into AI, but it's a nascent science. It's a nascent technology. We have many unanswered questions, including the socially relevant AI, including AI for good, including AI for education, healthcare, and many other areas.

So one of the biggest things I see would be investment into the basic science research into our universities and encouraging more students thinking in interdisciplinary terms, taking courses. You know, they can be STEM students, STEAM students. AI is not just for engineers and scientists; it could be for students with policymaking mind, for students with law interests, and so on. So I hope to see universities participating in this in a tremendous way, like many great schools in New York State.

Mr. TONKO. Thank you. Dr. Persons or Mr. Brockman, either of you?

Mr. BROCKMAN. Sorry. I agree with Dr. Li but I also point out that I think it is also becoming increasingly hard to truly compete as an academic institution because if you look at what's happening, industry right now is actually doing fundamental research. It's very different from most scientific fields in that the salary disparity between what you can get at one of these industrial labs versus what you can get in academia, it's a very, very large.

And there's a second piece, which is in order to do the research, you need access to massive computational resources. And, for example, the work that we just did with this, you know, game breakthrough that required basically a giant cluster of, you know, something around 10,000 machines. And that's something where in an academic setting it's not clear how you can get access to those resources. And so I think for the playing field to still be accessible, I think that there needs to be some story for how people in academic institutions can get access to that, and I think the question of, you know, where is the best research going to be done and where are the best people going to be, I think that's something that it's, you know, playing out right now in industry's favor, but it's not necessarily set in stone.

Mr. TONKO. Thank you. Dr. Persons?

Dr. PERSONS. Yes, sir. Thank you for the question. And I would just add to my fellow panelists the fact that our experts had said that real-world testbeds are important to this. You don't know what you don't know, so not only in addition to adding access to data but being able to test and do things, these—one thing for sure, and I learned, in fact, from OpenAI that a lot of the times these things come out with surprising results, and so that's the whole reason of creating safe environments to try things out and de-risk those technologies. And that's something that I think is going to be important to enable that base of research to have an avenue to perhaps move up the technology maturity scale possibly into the market and certainly hopefully to solve critical, complex, real-world problems.

Mr. TONKO. Thank you. Very informative. Mr. Chair, I yield back.

Mr. WEBER. The Chair now recognizes the gentleman from Illinois.

Mr. FOSTER. Thank you, Mr. Chairman. And thank you for coming to testify today.

You know, I've been interested in artificial intelligence for quite a long time. Back in the 1990s working in particle physics we were using neural network classifiers to have a look at trying to classify particle physics interactions. And when I couldn't stand it during the government shutdown and not so long ago, I went and downloaded TensorFlow and worked through the—part of the tutorial on it.

And, you know, the algorithms are not so different than what we were using back in the 1990s, but the computing power difference is breathtaking. And I very much resonated with your comments on the huge increase in dedicated computer power for deep learning and similar—and that is likely to be transformative, given the recent—and so that—you know, we have to think through that because even with no new brilliant ideas on algorithms, there's going to be a huge leap forward. So thank you for that. That's a key observation here.

You know, in Congress I'm the co-Chair of the New Democrats Coalition on the Future of Work Task Force where we have been trying to think through what this means for the workplace of the future. And so I'd like to—if—Mr. Chairman, I'd like to submit for the record a white paper entitled "Closing the Skills and Opportunity gaps," without objection.

Mr. WEBER. Without objection, so ordered.

[The information appears in Appendix II]

Mr. FOSTER. Thank you. And I will be asking for the record if you could have a look at this and see if—you know, how—what sort of coverage you think this document has for the near-term policy responses because it's—you know, this is coming at us I think faster than a lot of people in politics really understand.

And also, I will be asking for the record—I guess you may not have to respond right now—where the best sources of information on how quickly this will be coming at us. You know, there are conferences here and there, but you attend and your friends attend a lot of them. I'd be interested in where you think—you really come

together to get the techno-experts, the economic experts, you know, the labor economists, people like that all in the same room. I think it's something we should be putting more effort into.

On another track, I've been very involved in Congress in trying to resurrect something called the Office of Technology Assessment. You know, and what the J.O. did here is very good, which is to bring—we had a conference of the experts, and you brought in a good set of experts. And a year later now we are getting a report on this. And, you know, you need more bandwidth in Congress than that just, you know, on all technological issues, but this is a perfect example. A year-old group of experts and then AI, you know, is—those are opinions that are sort of dated a little bit, even a year in the past.

And so the Office of Technology Assessment for decades provided immediate high-bandwidth advice to Congress on all sorts of technological issues. And so we're coming closer and closer every year in getting it refunded after it was defunded in the 1990s. And so I think—well, to ask you a question here, is there anyone on the panel who thinks that Congress has enough technological capacity as it currently stands to deal with issues like this?

Dr. PERSONS. So——

Mr. WEBER. I can answer that.

Mr. FOSTER. Yes—no, it's a huge problem, and it's been aggravated by the fact that people have decided in their wisdom to cut back on the size and salaries available for Congressional staff. One of my—the previous members who talked about—here talked about the difficulty the federal government will have in getting real professionals, top-of-the-line professionals in here, and, you know, we're seeing Members of Congress who are willing to do anything but give them the salaries that will be necessary to actually compete for those jobs.

Let's see. I am now—let's see. Oh, Mr. Brockman, you had—your—I would advocate everyone have a look at the reference 5 in—which is your malicious use of AI, your—reference 5 in your testimony, which I spent—I stayed up way too late last night reading that, and it is real.

Along the same lines, Members of Congress have access to the classified version of a National Academies of Science study on the implication of autonomous drones for—and this is something that I think, you know, has to be understood by the military. We're about to mark up a military authorization bill, an appropriations bill that is spending way too much money fighting the last war and not enough fighting the wars of the future.

And then finally, Dr. Li, the—in the educational aspects of this, one thing I struggle for—I guess if you look through the bios of people who are the heroes of artificial intelligence, they tend to come from physics, math, places like that. And in theoretical physics or mathematics a huge fraction of the progress comes from a tiny fraction of people. It's just a historical truth. And I was wondering, is AI like that? Are they—you know, are there a small number of heroes that really do most of the work and everyone else sort of fills in the thing?

Dr. LI. So, like I said, Dr. Foster, AI is a very nascent field, so even though it is collecting a lot of enthusiasm worldwide, soci-

etally, as a science, it's still very young. And as a young science, it starts from a few people.

As a—I was also trained as a physics major, and I think about early days of Newtonian physics, and that was a smallish group of people as well. I mean, it's—it would be too much to compare directly, but what I really do want to say is that we—maybe in the early, even pre-Newtonian days of AI, we are still developing this, so the number of people are still small.

Having said that, there are many, many people who have contributed to AI. Their names might not have made it to the news, to the blogs, to the tweets, but these are the names that, as students and experts of this field, we remember them. And I want to say many of them are members of the underrepresented minority group. There are many women in the first generation of AI experts. So——

Mr. FOSTER. Yes. And when I was——

Dr. LI. —we need to hear more from them.

Mr. FOSTER. —you know, two or three clicks down in the references cited by your testimony and you look at the papers there and the author lists, it's pretty clear that our dominance in AI is due to immigrants, okay? And, Dr. Li, I suspect you might not have come to this country under the conditions that are now being proposed by our President. And I won't ask you to answer that, but it's important when we talk about what it is that makes this country dominant in things like AI. It is immigrants, okay? And I'll just leave it at that, and I guess my time is up.

Mr. WEBER. I thank the gentleman. I thank the witnesses for their testimony and the Members for their questions. The record will remain open for two weeks for additional written comments and written questions from Members.

The hearing is adjourned.

[Whereupon, at 12:24 p.m., the Subcommittees were adjourned.]

Appendix I

Answers to Post-Hearing Questions

ANSWERS TO POST-HEARING QUESTIONS

Responses by Dr. Jaime Carbonell

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University

Questions submitted by Chairwoman Barbara Comstock, House Committee on Science, Space, and Technology

1. Are the discoveries made in artificial narrow intelligence thus far reflective of a focus on certain types of artificial intelligence research? For example, is there a need for more human cognitive research to help facilitate development of artificial general intelligence (AGI)? What would be the most effective research path forward to advance AGI developments, and who should conduct such research – government, industry, academic institutions?

   Answer: First, the term "artificial narrow intelligence" is not commonly used by most researchers and practitioners in the field. Instead they call it "applied AI" or just "AI." Second, yes indeed, much of AI is focused on solving specific important problems such as self-driving vehicles, speech recognition, image recognition, special purpose robotics. There is also considerable work on broader aspects of intelligence such as natural language understanding, autonomous robots, underlying computational architectures such as deep neural networks, and so on. There is less work on trying to build a single unified intelligence. Many AI researches eschew such a goal in that they see AI as being a powerful suite of methods to improve just about everything rather than to substitute for humans. A few however pursue the (still distant) goal of true general intelligence. My belief and that of many, but far from all, AI researchers is that the best path towards AGI is largely the one we are on – solve first the difficult components of intelligent behavior, spin-off useful applications to keep the research funding whether based on the underlying science or the economic or military utility, or more typically a combination, and gradually build up to broader aspects of intelligence. This approach also gives us ample time to weigh the ethical implications as we evaluate the emerging AGI systems.

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University

Questions submitted by Representative Neal Dunn, House Committee on Science, Space, and Technology

1. With what confidence can we guarantee that we can prevent A.I.s that are psychopaths, such as MIT's Norman?

   Answer: Norman is a publicity stunt by a few members of the MIT Media Lab. MIT of course conducts much top-quality research, so my criticism is not leveled at the institution (my alma mater) but at this specific project. I do not know the inner workings of Norman, but from the information they provide, it appears to be based on machine learning algorithms being purposely fed false and misleading data. The result of course is a confused AI system. Humans, especially children, can be driven crazy, indoctrinated or brainwashed if exposed to volumes of false and misleading data with no true or accurate data to counteract it. The same is true in a more rudimentary way for machine learning systems. So there is no possible guarantee of sanity other than exposure to information that is mostly true and balanced, and not meant to deceive. An interesting line of new research in AI is to help identify fake data or postings, but that requires exposure to large amounts of normative (mostly correct) data for contrast. This line of research, identifying false reports, is particularly timely in light of Russian trolling and meddling in US elections, but still has quite a way to go.

2. Will we be able to teach them to act like they have compassion or empathy? What compels them to act this way?

   Answer: Yes, if we categorize behaviors as demonstrating degrees of compassion or empathy, then train machine learning models on contextually selecting said behaviors for appropriate circumstances in which to exhibit them – e.g. knowing when the AI system should say it is sorry, or when to say "better luck next time" or when to recall and retell a similar situation that caused concern, is definitely possible by extending current methods. However these approaches simulate empathy or compassion, even if done in contextually appropriate ways, and that is not the same as truly feeling such emotions. No one yet knows how or whether the latter may be doable.

3. Will we, as flawed humans, be able to create a perfect consciousness? What will the ramifications of that be?

   Answer: I will have to pass on this question, because I do not know what constitutes a perfect consciousness, and likely no one else really knows either.

4. Will we create an algorithm that will override data input?

   Answer: Presuming the question refers to data-driven machine learning algorithms, this is very much an active area of research. Isaac Asimov, the late well-known science fiction writer, got it right with his three laws of Robotics, wherein his imaginary future robots learned, adapted, and behaved in complex ways, but always obeyed the three laws, overriding input data when necessary to do so. My own research includes this question – how to create flexible neural-network learning systems that obey constraints (rules) even as they adapt based on input data. Neural nets without added mechanisms will forget earlier information if given overwhelming new data – the real question is how to not forget and always respect important constraints, and thus override data if necessary, but that requires creating additional structure for these algorithms. So the short form of the answer is "yes."

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University

Questions submitted by Representative Bill Foster, House Committee on Science, Space, and Technology

1. Your testimony mentioned China's goal of dominating the AI industry and that it has already surpassed the US in terms of patents granted for AI technologies. Do you have any thoughts on whether our current standard for what is eligible for patent protection helps or hinders AI innovation and American leadership in this space?

   Answer: I am not an expert on the legal framework for patentability, but the 2014 Supreme Court decision on Alice Corp. v. CLS Bank International dampened the patentability of software based methods and algorithms including AI methods. However more recent interpretations of the "Alice" decision have been better balanced. Some of my colleagues opine that patenting AI methods is more difficult these days than in pre-Alice times.

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Jaime Carbonell, Director, Language Technologies Institute, and Allen Newell Professor, School of Computer Science, Carnegie Mellon University

Questions submitted by Representative Jacky Rosen, House Committee on Science, Space, and Technology

1. Research grants and cooperative research agreements are the most common ways we support science and technology research at major universities and federal labs. In my district, the University of Nevada Las Vegas is working on various applications of artificial intelligence in areas including: health care, transportation, big data, and cyber security. Currently, through an NSF grant, UNLV is partnering with the local Clark County School District to mentor high school teachers on computer science, cybersecurity, and Big Data. This program will equip teachers to introduce Big Data and AI curriculum to high school students and hopefully inspire them to pursue further computing education and related careers. Across the country, we are continuing to see a huge demand for workers in STEM fields, but despite these increasing opportunities, not enough Americans have the necessary STEM education and skills.

    a. What more can we do to prepare and train enough individuals with the essential skills to meet the needs of an AI-driven economy today and in the future?

    Answer: This is a very timely question. As a country we must ramp up the training of AI-practitioners as well as AI researchers. The demand far exceeds the supply, just ask any hiring manager in the AI industry. Part of the problem is that universities have great difficulty retaining AI faculty, who are being offered outsize compensation packages by industry, thereby reducing the pool of AI educators. The issue is not just salary but research funding and facilities – and there the Government can help with increased NSF and other funding, to entice the cream of the crop to remain in academia and train the next generations of AI engineers and scientists. At CMU we recently created an AI major – that is AI at the undergraduate level, to supplement AI programs at the MS and PhD levels. But we too suffer challenges retaining our best educators. The creation of other such AI programs at other universities would be very useful, if they can hire and retain the appropriate faculty in today's very tight market. So, in short, the responsibility is shared between academia: create better advanced and undergraduate training in AI; and Government: create more research funding opportunities to both further the science and to retain the top academic talent in positions where they prepare the next generations of AI experts.

2. Research has shown that AI has the potential to reinforce existing biases. We need to ensure that the datasets used to input information about our society represent the diversity of the world we live in and address and prevent algorithmic bias.

   a. How can we ensure more diversity in datasets that feed into AI systems?

      Answer: Indeed, the question is timely. Biased data leads to biased results – this is true across science, and in AI it may also lead to biased AI systems. It is our burden to understand the biases, and when appropriate to counteract them. For instance, an AI system trained only on crime reports might conclude that most activities are criminal. If the purpose is to analyze patterns when crime is known to occur, then that bias is not harmful, as the results only apply to criminal activity. But if the purpose is to determine whether crimes occurred then the bias will be very harmful – the system would need a diverse and representative example of non-criminal activities as well. Hence coping with bias is more a burden on the researchers and AI engineers to either provide unbiased data or to interpret the results in light of the bias and only apply said results when the same conditions exist as were present for the data collection. The old adage "garbage in, garbage out" applies universally; my version is "bias in, bias out." AI systems are not immune.

   b. Could the fact that we still have a significant underrepresentation of women and minorities in the fields that feed into AI – such as computer science, physics, and engineering – contribute to this bias problem?

      Answer: I believe that underrepresentation is not the central problem in this specific data-bias problem. The central issue with training machine learning algorithms at the heart of AI systems is bias – deliberate or accidental – in the input training data. Any good statistician can vouch for that observation. The gender, race or ethnicity of the researcher is not directly a data-biasing factor. That said, however, the selection of which problems to address with powerful AI techniques can be more reflective of the researcher's background. Problems of concern to minorities may not be selected for study by non-minority researchers. And that is a real bias, which would be, to a significant extent, overcome by greater gender, race, ethnicity, and cultural balance among AI researchers.
      It is indeed the case that women and minorities are significantly underrepresented in the fields that feed AI graduate programs, and those gaps should be narrowed and eventually closed for many good reasons beyond data bias or problem-selection bias. I am proud to say that Carnegie Mellon University has achieved close to gender parity in its undergraduate computer science student population. Other universities are striving to reach that goal, but there is a quite a way to go.

And we must all work harder to be more inclusive of minority students in these fields. My own personal experience having advised some eight PhD students, over two dozen of whom being women, and a small handful of minority PhD students, is that role models and peers play a strong role in their success. Being of Hispanic descent has helped me better understand their perspectives and concerns, as they are similar to what mine were at the time I studied physics and mathematics on the road to becoming an AI researcher.

Responses by Dr. Tim Persons

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office

Questions submitted by Chairwoman Barbara Comstock, House Committee on Science, Space, and Technology

1. Are the discoveries made in artificial narrow intelligence thus far reflective of a focus on certain types of artificial intelligence research? For example, is there a need for more human cognitive research to help facilitate development of artificial general intelligence (AGI)? What would be the most effective research path forward to advance AGI developments, and who should conduct such research – government, industry, academic institutions?

   Answer: Our technology assessment on artificial intelligence (AI) did not explicitly focus on narrow AI versus AGI. However, our assessment does describe how AI has been defined in a variety of ways, an approach for conceptualizing the development of narrow AI and AGI, and policy and research areas that should be addressed going forward.

   Specifically, we highlighted several policy issues that the experts who participated in our 2017 forum on AI believe require further attention, including:
   - incentivizing data sharing,
   - improving safety and security,
   - updating the regulatory approach, and
   - assessing acceptable risks and ethical decisions.

   As indicated by the range of forum participants, research should be conducted by government, industry, and academics as all have a role to play in this emerging area.

2. What are some ways in which individuals or nation states might use artificial general intelligence (AGI) maliciously in a physical or cyber-attack? Could AGI cyber-attacks wreak more havoc and damage than the types of major cybersecurity hacks reported in recent history such as the Office of Personnel Management data breach, Equifax data breach, Yahoo data breach, etc.?

   Answer: While we have not done work to compare potential damages caused by AGI versus other types of cyber attacks, participants from our forum highlighted challenges and opportunities to enhancing the safety and security of system applications from cyber attacks, including those with AI features. One participant emphasized that security of the system is only as good as the security of the computer that it is running on, so if an AI

system is running on a computer that is then hacked, the system itself is also vulnerable. In addition, some participants stated that protecting against hackers is not something an individual company can do on its own. Rather, efforts to combat hackers need to be industry-wide.

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office

Questions submitted by Representative Neal Dunn, House Committee on Science, Space, and Technology

1. With what confidence can we guarantee that we can prevent A.I.s that are psychopaths, such as MIT's Norman?

   Answer: While our technology assessment on artificial intelligence did not explicitly address this issue, our assessment does highlight concerns, policy implications, and research areas to consider related to ethical decision making in the context of AI.

2. Will we be able to teach them to act like they have compassion or empathy? What compels them to act this way?

   Answer: Our technology assessment on artificial intelligence did not address how closely AI will be able to learn compassion and empathy, if ever. However, some of the experts who participated in our 2017 forum on AI highlighted the need to develop an ethical framework to govern the use of AI and ensure the actions and decisions of AI systems can be adequately explained and accepted by those who interact with such systems. For example, some of the participants noted that before humans will understand, appropriately trust, and be able to effectively manage AI, an AI application or system needs to explain why it took certain actions and why it valued certain variables more than others. Accordingly, a key policy consideration will be to assess acceptable levels of risk and ethical considerations of AI systems.

3. Will we, as flawed humans, be able to create a perfect consciousness? What will the ramifications of that be?

   Answer: This is another area that could be considered in terms of developing an ethical framework. Forum participants highlighted the need for research in exploring computational ethics and explainable AI, whereby systems can reason without being told explicitly what to do and inspect why they did something, making adjustments for the future. Participants noted the need to develop and adopt an appropriate ethical framework to govern the use of AI in research, as well as explore factors that govern how quickly society will accept AI systems in their daily lives.

4. Will we create an algorithm that will override data input?

   Answer: Our technology assessment did not specifically address this issue. This is another issue that would have to be considered as part of developing an ethical framework as this area continues to develop.

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office

Questions submitted by Representative Bill Foster, House Committee on Science, Space, and Technology

1. Your testimony briefly touched on a forum participant's concerns about the adequacy of current laws to protect products created by AI. Do you have any thoughts on whether our current standard for what is eligible for patent protection helps or hinders AI innovation and American leadership in this space?

   Answer: In our technology assessment, we pointed out that the widespread adoption of AI raises questions about the adequacy of current laws and regulations. For example, deploying AI requires contractual agreements with the users of the AI, according to one participant from our 2017 forum on AI, and contracts fall under state law, which could be an issue when state laws create rights through contracts that are not protected by the federal government.

   In another example, a participant noted that current patent and copyright laws provide only limited protection for software and business methods and questioned whether these laws will adequately protect the products created by AI. This same participant also was concerned that current patent and copyright laws pose challenges for AI in at least three key areas:

   - An inventor does not have to reveal very much about their software code to secure a patent or copyright, which may be problematic where public safety, liability, or criminal justice is concerned.

   - Patent protection lasts 20 years and copyrights created by corporations last 120 years, both of which are considered too long a time horizon for AI, according to this participant. Specifically, because advancements in AI have moved at such an extraordinary pace, this participant believes that data protection for pharmaceuticals may be a better model for protecting innovations in AI. In that context, a brand-name drug company receives 4-5 years of exclusivity in exchange for making its safety and efficacy clinical trial data available to potential competitors. Thus, pharmaceutical companies receive this period of exclusivity, enforced through the context of regulatory approval, in exchange for data openness.

- If AI derives its creative results in part through the collective actions of numerous humans, according to this participant it is not clear whether that creativity is attributable solely to the program or to the program's creators.

In our technology assessment, we discussed concerns and potential approaches for a government role in mitigating such concerns. For instance, participants emphasized the need for establishing a "safe space" to protect sensitive information (e.g., intellectual property, privacy, and brand information). Another participant reiterated optimism that government could get the data it needed to properly protect the public, while maintaining proprietary data protections. In this regard, participants noted successful data-sharing efforts through entities such as MITRE and the National Institute of Standards and Technology (NIST).

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Tim Persons, Chief Scientist, U.S. Government Accountability Office

Questions submitted by Representative Jacky Rosen, House Committee on Science, Space, and Technology

1. Research grants and cooperative research agreements are the most common ways we support science and technology research at major universities and federal labs. In my district, the University of Nevada Las Vegas is working on various applications of artificial intelligence in areas including: health care, transportation, big data, and cyber security. Currently, through an NSF grant, UNLV is partnering with the local Clark County School District to mentor high school teachers on computer science, cybersecurity, and Big Data. This program will equip teachers to introduce Big Data and AI curriculum to high school students and hopefully inspire them to pursue further computing education and related careers. Across the country, we are continuing to see a huge demand for workers in STEM fields, but despite these increasing opportunities, not enough Americans have the necessary STEM education and skills.

    a. What more can we do to prepare and train enough individuals with the essential skills to meet the needs of an AI-driven economy today and in the future?

    Answer: Experts who participated in our 2017 forum on AI noted the need to have research to reevaluate and reimagine training and education as well as better understand the implications of AI on employment. Additionally, a participant suggested that further research is needed to explore new means to encourage students from low-income backgrounds, women, and minorities to pursue science, technology, engineering, and math (STEM) subjects, where the majority of jobs will be going into the 21st century. In recent prior work, we found that certain minority groups made gains in representation in the technology workforce from 2005 to 2015, but that no growth occurred for other minority groups.[1] Researchers interviewed for that work identified several factors that may have contributed to lack of growth for some groups, including fewer women and minorities graduating with technical degrees and company hiring and retention practices.

    In addition to R&D grants, each year the federal government manages a portfolio of grant programs to promote STEM education. In our most recent work in this

[1]GAO, *Diversity in the Technology Sector: Federal Agencies Could Improve Oversight of Equal Employment Opportunity Requirements*, GAO-18-69, (Washington, D.C.: Nov. 16, 2017).

area, we noted that the interagency body responsible for implementing the federal STEM education strategic plan has not reviewed programs' performance assessments or reported the participation rates of underrepresented groups in these programs, as required by law.[2] Reviewing and reporting this information is important for assessing whether programs are broadening access to groups historically underrepresented in STEM fields.

In our forum, we also highlighted the importance of bringing those with STEM skills together with those from other disciplines. For instance, one participant emphasized that students who want to work with AI will need to master STEM disciplines, as well as the social sciences, because the technology that is developed is going to interact with humans.

2. Research has shown that AI has the potential to reinforce existing biases. We need to ensure that the datasets used to input information about our society represent the diversity of the world we live in and address and prevent algorithmic bias.

   a. How can we ensure more diversity in datasets that feed into AI systems?

      Answer: Although our assessment did not directly address this issue, forum participants emphasized that if the data being used by AI are biased, the results will be biased as well. One participant noted that with AI, there is no guarantee of freedom from bias. AI can help prevent inappropriate or harmful human bias, according to this same participant, if it is carefully used, if the assumptions of the models are thoughtfully considered, and, most importantly, if the outputs of the model are constantly and closely verified. Another participant stated that the baseline is current practice, not perfection, and that the goal should be to become less biased and more accurate, not perfect.

   b. Could the fact that we still have a significant underrepresentation of women and minorities in the fields that feed into AI – such as computer science, physics, and engineering – contribute to this bias problem?

      Answer: This is not an issue that we have looked at directly as it relates to AI. However, in a report we issued on women representation in STEM research,[3] we noted that research has found that diversity brings different information, opinions, and perspectives to a group, leading to better decision-making than non-diverse groups.[4]

---

[2]GAO, *Science, Technology, Engineering, and Mathematics Education: Actions Needed to Better Assess the Federal Investment*, GAO-18-290, (Washington, D.C.: March 23, 2018).
[3]GAO, *Women in STEM Research: Better Data and Information Sharing Could Improve Oversight of Federal Grant-making and Title IX Compliance*, GAO-16-14, (Washington, D.C.: Dec. 3, 2015).

[4]Moss-Racusin, Corinne A. et al. "Science faculty's subtle gender biases favor male students" Proceedings of the National Academy of Sciences 109, no. 41 (2012).

Responses by Mr. Greg Brockman

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Mr. Greg Brockman, Co-Founder and Chief Technology Officer, OpenAI

Questions submitted by Chairwoman Barbara Comstock, House Committee on Science, Space, and Technology

1. Are the discoveries made in artificial narrow intelligence thus far reflective of a focus on certain types of artificial intelligence research? For example, is there a need for more human cognitive research to help facilitate development of artificial general intelligence (AGI)?

   Answer: Most AI research today involves a machine learning technique called deep learning. As compared to previous techniques, deep learning is more focused on empirical results than theory.

   That said, AI researchers are constantly combining deep learning techniques with techniques from other scientific disciplines, including neuroscience and graph theory. We at OpenAI — and AI researchers everywhere — still probably have many blind spots about how best to build safe and broadly beneficial AGI. In the coming years, I expect many AGI-relevant insights to come out of scientific disciplines other than machine learning.

   What would be the most effective research path forward to advance AGI developments, and who should conduct such research – government, industry, academic institutions?

   Answer: This question can be considered from the perspective of research effectiveness, and from the perspective of values or incentives.

   From the research-effectiveness perspective: Progress toward AGI benefits from less-conventional institutional settings. Progress requires both scientific progress — i.e., rapidly testing hypotheses and running experiments — and large-scale engineering efforts. We've deliberately structured OpenAI to strike this balance between resources and nimbleness. As a result, we've been able to rapidly iterate — i.e., carry out a large number of hypothesis→experiment→result cycles — while simultaneously using substantial amounts of computing infrastructure.

   Properties of more-conventional institutional settings are, nonetheless, critical to AI progress. Specifically, progress toward AGI will likely require the intellectual

environment of academic centers, the resources of industry, and the flexibility and experimental mentality that is the hallmark of a Silicon Valley startup.

From the values/incentives perspective: It's important that any institutions developing advanced AI technology are structurally aligned with the interests of humanity. That is to say: they are committed to the broad distribution of benefits from AI, and they have a safety-first mindset for developing AI. OpenAI is structured to serve this function, but in principle, other institutions, with different structures, could serve this function.

2. What are some ways in which individuals or nation states might use artificial general intelligence (AGI) maliciously in a physical or cyber-attack? Could AGI cyber-attacks wreak more havoc and damage than the types of major cybersecurity hacks reported in recent history such as the Office of Personnel Management data breach, Equifax data breach, Yahoo data breach, etc.?

   Answer: I'll refer in this answer to the risks of "advanced AI" more broadly, which are on a spectrum of generality and capability beyond today's AI systems. Advanced AI, like all technologies, can be used to amplify our (human) ability to achieve our goals, whatever those goals might be. Hopefully, the goals most of us want to pursue are generally agreeable — e.g., be more economically productive.

   Unfortunately, some people have less-agreeable goals, and — absent social or technical safeguards — advanced AI will empower them as well. Using advanced AI to augment illegal or unjustified physical- or cyber-attacks is just one such example. Advanced AI might not only increase the risk of these types of attacks; it might also make it more difficult to identify the source of the attack, and lower the barriers to entry for attacks.

   We recently co-authored a report on the risks of the malicious use of AI, which goes into substantial detail about actions that can be taken today to address these risks, including, for example, research into automated exploit generation.[1]

[1] https://arxiv.org/pdf/1802.07228.pdf

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Mr. Greg Brockman, Co-Founder and Chief Technology Officer, OpenAI

Questions submitted by Representative Neal Dunn, House Committee on Science, Space, and Technology

1. With what confidence can we guarantee that we can prevent A.I.s that are psychopaths, such as MIT's Norman?

   Answer: Antisocial or unsafe behavior from an AI can be a result of either deliberate or inadvertent training. Both of these safety issues need to be addressed, but in different ways.

   MIT Media Lab's "Norman" experiment was a deliberate attempt to highlight how sensitive the characteristics of an AI system are to the data used to train it.[2] The MIT researchers chose to train an AI system on a variety of frightening images; they then asked the trained system to label new images that were not part of its training dataset. The system selected many frightening or disquieting labels for these images precisely because of how it was trained. This experiment highlights an important risk, but it's critical not to misinterpret the experiment as evidence that psychopath-like properties might be emergent in AI systems.

   That said, there are also *unintentional* safety risks related to the training and reward structure of AI systems, such as bias in an algorithm used in hiring or parole decisions. MIT's Media Lab is working on solutions to these risks. So are, for example, Microsoft's Fairness, Accountability, Transparency, and Ethics in AI ("FATE") group,[3] and Google Research.[4]

   OpenAI has a team devoted to working on how best to avoid a subclass of these risks, specific to more-advanced AI systems. For example, see our collaboration with DeepMind on "Learning from Human Preferences,"[5] which explores a training technique that might become more critical as AI systems become more advanced.

---

[2] https://www.media.mit.edu/projects/norman/overview/
[3] https://www.microsoft.com/en-us/research/group/fate/
[4] See, e.g., Mitchell, "Algorithmic Bias in Artificial Intelligence: The Seen and Unseen Factors Influencing Machine Perception of Images and Language," https://www.clsp.jhu.edu/events/margaret-mitchell-google-research/#.W1_Zu9hKiL9
[5] https://blog.openai.com/deep-reinforcement-learning-from-human-preferences/

2. Will we be able to teach them to act like they have compassion or empathy? What compels them to act this way?

   Answer: AI systems' "values" are a function of their learning procedure and training environment. AI systems acquire values implicitly, similar to the way humans learn values: they develop internal biases and subjective judgments, which reflect the data they are exposed to during training. This is cause for both optimism and concern. It's cause for optimism in that it means we should, in principle, be able to train AI systems to display the types of compassion and empathy that we ourselves display. It's cause for concern to the extent the level of empathy and compassion among humans today is lower than we would like it to be.

   As noted above, this is an area of active research at OpenAI.

3. Will we, as flawed humans, be able to create a perfect consciousness? What will the ramifications of that be?

   Answer: This is a complicated question that could easily be misinterpreted. To accurately answer it would likely require further clarification and discussion.

   With that caveat: as is the case for any technology, it's important that we build safe AI systems, the behavior of which is aligned with human goals, including human values. OpenAI works toward this goal in several respects. For example, it does fundamental research related to developing advanced AI systems that can be aligned with human goals. It also strives to cultivate a broader interest in the machine learning community in safe and socially responsible AI research and development.

   For a high-level overview of OpenAI's commitment to these principles, see its Charter.[6]

4. Will we create an algorithm that will override data input?

   Answer: AI systems' actions are the result of their learning procedure and training environment. Sometimes, the actions a system takes in pursuit of an apparently straightforward objective can be concerning or surprising.

   For example, we once trained an AI system to play a simulated boat race game by trying to maximize the game score.[7] We assumed the system would learn to complete the race quickly, believing that was the best way to achieve a high score. But the system found another way: It found a lagoon where it could turn in a large circle and repeatedly knock

[6] https://blog.openai.com/openai-charter/
[7] https://blog.openai.com/faulty-reward-functions/

over targets, timing its movement so as to always knock over the targets just as they repopulate, while catching on fire and crashing into other boats.

The boat race experiment is an example of an AI system achieving an apparent straightforward goal in a surprising — and potentially dangerous — way. The problem is that the *actual* goal of the human designers was not simply to maximize game score, but to do so in a way consistent with the spirit of the game. But programming the "spirit of the game" as part of the program's formal objective is much more complicated. This example evidences a potentially concerning safety risk, but it's important to distinguish this risk from the risk of an algorithm overriding data input. The algorithm did not ignore the data; rather, it found a new, unexpected way of achieving a particular outcome.

We've co-authored a paper, which describes several technical problems AI researchers will need to solve to ensure AI systems act predictably and in accordance with our *actual* objectives.[8]

[8] https://arxiv.org/abs/1606.06565

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Mr. Greg Brockman, Co-Founder and Chief Technology Officer, OpenAI

Questions submitted by Representative Jacky Rosen, House Committee on Science, Space, and Technology

1. Research grants and cooperative research agreements are the most common ways we support science and technology research at major universities and federal labs. In my district, the University of Nevada Las Vegas is working on various applications of artificial intelligence in areas including: health care, transportation, big data, and cyber security. Currently, through an NSF grant, UNLV is partnering with the local Clark County School District to mentor high school teachers on computer science, cybersecurity, and Big Data. This program will equip teachers to introduce Big Data and AI curriculum to high school students and hopefully inspire them to pursue further computing education and related careers. Across the country, we are continuing to see a huge demand for workers in STEM fields, but despite these increasing opportunities, not enough Americans have the necessary STEM education and skills.

    a. What more can we do to prepare and train enough individuals with the essential skills to meet the needs of an AI-driven economy today and in the future?

    Answer: One way to encourage more people to learn AI-relevant skills is to recognize and reward excellence in AI research — e.g., through prizes, awards, fellowships, or special programs. For example, the vanguards of early space development were highly regarded by their communities, which in turn inspired a generation of children to dream of becoming astronauts. Similarly, the physicists of the mid-1900s received great renown and prestige, leading to many of the brightest minds of a generation to learn physics. I hope that we can inspire a new generation of AI researchers in a similar way.

    A more-direct incentive than prestige is funding. Funding often comes with conditions that make it less attractive to researchers. Increasing the amount of no-strings-attached grants for AI research --- particularly pro-social research that might be insufficiently incentivized in commercial markets -- could be highly effective. A related mechanism is direct funding of PhD students or PhD programs.

    But prestige and funding are insufficient if the best minds of the next generation don't have the requisite skills. To address that risk, we should invest more in

STEM education. These investments should target both early-age educational programs, as well as more-practical vocational programs for older students and adults.

2. Research has shown that AI has the potential to reinforce existing biases. We need to ensure that the datasets used to input information about our society represent the diversity of the world we live in and address and prevent algorithmic bias.

   a. How can we ensure more diversity in datasets that feed into AI systems?

   Answer: We should work toward a research and industry culture that normalizes testing for bias, fairness, and interpretability — among other risks — throughout the research, development, and deployment process. In practice, this takes many forms:

   First, we should make sure that the datasets we use to train AI systems reflect the diversity of their real-world environments. For example, a recent Washington Post analysis showed how widely deployed smart speakers from Google and Amazon (Google Home and Alexa) favor certain accents, likely as a consequence of being trained on data that favors those accents.[9]

   Second, we should bear in mind historical human biases that might be deeply ingrained in the data on which AI systems are trained.[10] If our goal is to train AI systems that can improve the world, we should take care not to entrench the biases of the past.

   Third, we should seek international approaches to international problems, both as a means of obtaining more-diversified datasets and as means of achieving positive-sum solutions to global challenges.

   Lastly, governments might consider creating — or catalyzing the creation of — datasets that are more representative of the entire population the governments represent. If such datasets existed, private organizations — such as OpenAI — wouldn't depend so heavily on the selective, market-driven datasets developed in the commercial sector.

[9] https://www.washingtonpost.com/graphics/2018/business/alexa-does-not-understand-your-accent/?noredirect=on&utm_term=.32f0e25b4996

[10] See, e.g., Crawford, "The Trouble with Bias," https://nips.cc/Conferences/2017/Schedule?showEvent=8742

b. Could the fact that we still have a significant underrepresentation of women and minorities in the fields that feed into AI – such as computer science, physics, and engineering – contribute to this bias problem?

Answer: Yes. AI is both a technology that can transform its environment, and a reflection of the environment in which it's created. As I note above, this fact creates a risk of entrenching current cultural biases in our technology. I cite above several current research efforts — at many different organizations — working to solve these problems. OpenAI, for its part, has started an OpenAI Scholars program to create paid mentorship for high-potential people from backgrounds underrepresented in the field.[11] But much more is happening than I have cited, and much more needs to happen still.

[11] https://blog.openai.com/openai-scholars/

Responses by Dr. Fei-Fei Li

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Fei-Fei Li, Chairperson of the Board and Co-Founder, AI4ALL

Questions submitted by Chairwoman Barbara Comstock, House Committee on Science, Space, and Technology

1. Are the discoveries made in artificial narrow intelligence thus far reflective of a focus on certain types of artificial intelligence research? For example, is there a need for more human cognitive research to help facilitate development of artificial general intelligence (AGI)? What would be the most effective research path forward to advance AGI developments, and who should conduct such research – government, industry, academic institutions?

   Answer: I'm a strong proponent of a cross-disciplinary approach to AI, in which complementary but distinct fields play a collaborative role. Since its inception, our field has benefited from the insights yielded by research in fields like neuroscience and cognitive science, and this trend should be encouraged as the technology develops. For example, today's Deep Learning algorithms have benefited greatly from Nobel Prize winning research in the mammalian primary visual cortex from the 1960's.

   Much of this collaboration will begin at universities, of course, as so much of the requisite talent is already sharing a campus, but our attitude towards the traditional separation of disciplines will have to be modernized in order to foster the level of cross-pollination the future of this technology will demand. Industry players will have to recognize that the creation of truly beneficial machine intelligence is not merely an engineering problem, and be willing to augment their own development teams with expertise that lies beyond computer science. Finally, governments can use funding, public research projects and even legislation to ensure this work remains transparent and committed to the common good.

   I caution against the notion of "AGI" as the focus of this endeavor, however. Across the field, little consensus exists on whether a distinction between "AGI" and "AI" is necessary, or even whether it makes sense. More importantly, AI is *already* playing a pivotal role in our world, and its influence is growing apace. This impact is being felt in finance, the criminal justice system, and even healthcare, and developments already on the horizon promise a transformative effect on the lives of people all over the world. By basing the AI of today *and* tomorrow on humanistic, cognitive, and even empathic insights from across the academic spectrum, we can ensure its impact is a positive one.

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Fei-Fei Li, Chairperson of the Board and Co-Founder, AI4ALL

Questions submitted by Representative Neal Dunn, House Committee on Science, Space, and Technology

1. With what confidence can we guarantee that we can prevent A.I.s that are psychopaths, such as MIT's Norman?

   Answer: It's important to remember that AI is still a nascent field, and that many important questions have yet to be answered. This makes it difficult to talk in terms of "confidence", especially on such speculative topics. However, MIT's Norman demonstrates the remarkable susceptibility of today's AI to the data and experiences from which it learns. If we want our algorithms to treat humans with respect and compassion, it's up to us to prioritize such values when developing them. In fact, I often end talks by reminding the audience that the behavior of machines will always be our responsibility, no matter how autonomous they become.

2. Will we be able to teach them to act like they have compassion or empathy? What compels them to act this way?

3. Will we, as flawed humans, be able to create a perfect consciousness? What will the ramifications of that be?

4. Will we create an algorithm that will override data input?

   Answer to 2, 3, and 4: As I mentioned above, we're so early in the evolution of this technology that we should approach any speculation with caution, especially when such complex ideas are involved. Our understanding of consciousness remains both rudimentary and contentious, with ongoing debates from philosophy to neuroscience. However, whether or not consciousness will someday play a role in AI, we know that today's machines can already reflect our worst traits—and even amplify them. In light of this, I believe we have an obligation to learn all we can from such mistakes, contain them before they grow, and take intelligent steps to help avoid them in the future.

   I liken this to the scientific method itself. Human perception is inherently limited—and easily misled, especially at very small and very large scales—and the scientific revolution was perhaps the first widespread acknowledgement of this awkward truth. Fortunately, our blind spots didn't preclude our ability to recognize them, and even build a system of reasoning to help us work around them. Likewise, although we're far from ethically perfect beings, I'm optimistic we can develop practices that anticipate those imperfections in the technology we create.

We're certain to make mistakes in the coming years. Whether or not we respond to them productively will be determined by our ability to talk honestly about them. Our greatest asset will be our humility.

HOUSE COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

"Artificial Intelligence – With Great Power Comes Great Responsibility"

Dr. Fei-Fei Li, Chairperson of the Board and Co-Founder, AI4ALL

Questions submitted by Representative Jacky Rosen, House Committee on Science, Space, and Technology

1. Research grants and cooperative research agreements are the most common ways we support science and technology research at major universities and federal labs. In my district, the University of Nevada Las Vegas is working on various applications of artificial intelligence in areas including: health care, transportation, big data, and cyber security. Currently, through an NSF grant, UNLV is partnering with the local Clark County School District to mentor high school teachers on computer science, cybersecurity, and Big Data. This program will equip teachers to introduce Big Data and AI curriculum to high school students and hopefully inspire them to pursue further computing education and related careers. Across the country, we are continuing to see a huge demand for workers in STEM fields, but despite these increasing opportunities, not enough Americans have the necessary STEM education and skills.

    a. What more can we do to prepare and train enough individuals with the essential skills to meet the needs of an AI-driven economy today and in the future?

    Answer: In addition to encouraging STEM education across the board, more can be done to open these fields to traditionally underrepresented groups like girls and people of color. For example, my own non-profit, AI4ALL, runs a growing range of summer programs across the country in which high school students from underserved communities get hands-on exposure to AI research in real university labs. Our own polling suggests this has a transformative effect on what's possible in the minds of such students, and I strongly advise we expand this effort.

    However, while diversity is a worthwhile goal by any standard—especially in a country founded on the value of inclusion—there's an argument to be made on purely numerical grounds as well: the future will soon present challenges that can only be met with a wide, deep pool of talent, and we simply can't afford to turn away qualified contributors.

2. Research has shown that AI has the potential to reinforce existing biases. We need to ensure that the datasets used to input information about our society represent the diversity of the world we live in and address and prevent algorithmic bias.

    a. How can we ensure more diversity in datasets that feed into AI systems?

b. Could the fact that we still have a significant underrepresentation of women and minorities in the fields that feed into AI – such as computer science, physics, and engineering – contribute to this bias problem?

Answer to both a and b: Bias in data can be addressed in two primary ways. First, the process by which data is organized must be made more sensitive to our myriad blind spots as humans, and more adept at counteracting them. This is an area of active research, and while there's a long way to go, I'm happy to say we're seeing early signs of its adoption in industry.

Second, and equally importantly, the people who take part in this process must *themselves* reflect the diversity of the world at large. Despite growing efforts to increase representation in tech, it remains predominantly white and male. Addressing this disparity will go a long way towards ensuring our technology treats everyone with fairness. This is another reason why organizations like AI4ALL are so important; more diversity in the AI industry means better experiences for its users, far beyond Silicon Valley.

Appendix II

Additional Material for the Record

OPENING STATEMENT SUBMITTED BY DR. CARBONELL

Testimony of Professor Jaime Carbonell

before the Committee on Science, Space, and Technology, Subcommittee on Research and Technology and Subcommittee on Energy, of the U.S. House of Representatives on the hearing titled,

"Artificial Intelligence - With Great Power Comes Great Responsibility"

June 26, 2018

Good Morning. My name is Jaime Carbonell. I am a professor of computer science and the director of the Language Technologies Institute at Carnegie Mellon University.

I would like to thank Chairman Comstock of the Research and Technology Subcommittee and Chairman Weber of the Energy Subcommittee for inviting me to testify today. I would also like to thank Ranking Member Lipinski, Ranking Member Veasey, and all of the Members of the Committee for your interest in Artificial Intelligence. It is a field in which I have been working for over 40 years.

Background: Artificial intelligence (AI) is the ability to create machines who perform tasks normally associated with human intelligence, from autonomous driving and spoken language comprehension to medical diagnosis and scientific discovery. Throughout its roughly 60-year history, AI has been incubated, admired, dismissed as infeasible, respected, maligned, feared, and more recently mainstreamed. Progress has been inexorable over the years, based on faster hardware, larger memories, new computing paradigms such as parallelism in graphical processing units (GPUs), major algorithmic innovations such as deep neural networks, and a plethora of successful applications driving the current industrial AI boom.

A Brief History: The term "Artificial Intelligence" was coined in 1956 at a conference in Dartmouth University by the four founders of the field: Herbert Simon, Alan Newell, Marvin Minsky, and John McCarthy, in their shared conviction that human thought could be reduced to symbolic computation on digital computers in the near future. Simon and Newell founded the AI efforts at Carnegie Mellon University, Minsky at MIT, and McCarthy at Stanford. Although there were some early successes, such as propositional logic provers and checker-playing programs, the true magnitude and complexity of human reasoning and the difficulty in replicating aspects of human intelligence in digital computers ensured that AI was a very long-term endeavor, not just a single moonshot or even a Manhattan Project.

The 1960's witnessed the founding of most of the AI subfields: computer vision, natural language processing, speech recognition, robotics, automated reasoning and planning, and so on. In the 1970's symbolic approaches to AI dominated, including reasoning based on first-order

logic, and in the 1980's the first substantial commercial benefits of AI were reaped with the advent of rule-based "expert-systems". These systems encoded narrow human expertise in commercially important areas such as automatically configuring digital computers, or deciphering mass-spectrograms. Another example of deep narrow expertise was in chess playing machines. IBM's Deep Blue, based on the earlier Deep Thought chess playing program from Carnegie Mellon University, beat world champion Gary Kasparov in 1989. Additionally, the field of modern Machine Learning was founded in that decade. During the 1990's steady progress was made on virtually all the component AI technologies, including neural networks for learning, combining vision, navigation and manipulation into operational robotics, and combining speech recognition and natural language processing into dialog systems. The 90's also heralded the statistical approaches in AI and machine learning, reducing many cognitive problems to mathematical optimization. Strangely, the 1990's are sometimes called the "AI winter", since much of the progress was not externally visible, and other technologies such as the advent of the internet and early electronic commerce captured the limelight.

After the turn of the century, industry started taking AI more seriously. The search engine companies, for instance worked on semantic associations to find information beyond the exact keywords in a query. For instance, asking for "inexpensive vehicle insurance" yields results about "cheap car insurance." Those same companies accessing huge amounts of electronic text in multiple languages to train cross-language associations were able to build the first widely-used general-purpose machine translation engines. The rule-based expert systems of earlier decades evolved and became embedded into manufacturing systems and workflow-planning, and otherwise entered routine practice, and were no longer the subject of academic inquiry. Instead, academia sought new challenges, such as integrated virtual agents, autonomous robots, multilingual speech-to-speech translation, and so forth. The movement of "Artificial General Intelligence" (AGI) took root where the focus on performing very well at a narrow difficult task was replaced by the desire to exhibit adaptive intelligence to perform reasonably across many tasks, i.e. re-focusing on the initial goal of AI to create human-like intelligence.

In the current decade interest in AI has exploded, in large part due to the resurgence of deep networks in machine learning, at the core of AI, and their ability to perform many tasks much better than before – tasks ranging from self-driving vehicles and robotic assistants to robust speech understanding to semi-intelligent chatbots such as Siri, Cortana and Alexa. All the large tech companies have sprouted AI labs or divisions, as have many in other disciplines including finance and manufacturing.

AI and its sub-disciplines: Borrowing the title of the famous 1955 movie we can say that human intelligence is a many-splendored thing; it is also a many-faceted thing. Vision and physical coordination is not a trait unique to humans, but inherited from our ancestor species. On the other hand, complex language, writing, higher mathematics and abstract reasoning, most scientists would attribute uniquely to our species or at least our genus. AI attempts to understand and replicate all aspects of human intelligence, including those that evolution took many millions

of years to perfect. But, to make the ultimate challenge somewhat more tractable, AI is divided into multiple somewhat overlapping disciplines:

- ***Sensing***: Using vision, sound, touch and super-human sensors such as sonar, radar, and lidar to sense and form a mental model of the immediate environment, and update the model dynamically.
- ***Communicating***: Using language, written or verbal, to convey information, questions or commands to other agents in the social group, as well as to record knowledge for future reference.
- ***Acting***: Effecting changes in the physical world by manipulating objects, navigating to other locations, observing the outcomes of the actions, and recording the results
- ***Reasoning***: Envisioning performing physical or communicative actions in the world, inferring their consequences, and planning sequences of actions to achieve desired goals, as well as inferring the likely actions of other agents.
- ***Learning***: Acquiring knowledge and experience to improve other cognitive tasks: sensing, acting, communicating and reasoning, including the ability to solve new problems or address new situations.

Many established AI areas map onto one or a combination of the above disciplines. For instance, computer vision is primarily sensing, but also learning to recognize, store and classify patterns. Natural language processing is primarily communicating, but contains elements of reasoning (e.g. semantic interpretation) and learning. Robotics, while centered on sensing and acting, touches on all the AI disciplines. Sometimes AI disciplines are confused with the overall endeavor, as in "are you doing AI or machine learning?" This is like asking "are you a car repairman or an engine mechanic"? If someone is practicing the latter they are also practicing the former. Similar confusion is evident in "should we invest in robotics or AI?" Or, "Should we build a chatbot rather than doing AI?" Instead the better question is "should we focus in Robotics – or on chatbots – as the most relevant part of AI for our business?" All of the subfields of AI are reinforced by advances in adjacent subfields, and some like machine learning are cross-cutting and ubiquitous.

Applications of AI: In many ways AI is the ultimate meta-technology, the technology that enhances most other technologies whether they be in strategic planning, manufacturing, healthcare, transportation, customer contact, entertainment, forensics, mining, agriculture, or even scientific research. AI extends to defense and national security, ranging from intelligence gathering and interpretation to operational warfighter capabilities. And, AI is playing an increasing role in humanitarian areas such as workplace safety and disaster relief. Let us look at a few illustrative applications of AI, focusing on the areas this writer knows best.

- ***Question Answering***: Wouldn't it be great if we could bring to bear all the world's knowledge – or at least all the publicly available knowledge – to bear in answering

automatically every burning question? Starting with DoD funding, followed by IBM's Watson in the Jeopardy Challenge, to today's open-domain systems, Q&A has emerged as a major challenge task for AI. I am proud to say that Carnegie Mellon has been a central player in Q&A from the get-go in government programs, through helping IBM build the original Watson, to heading the leader boards in current Q&A challenges.

- ***Autonomous Driving***: Imagine drivers who do not get distracted, do not get sleepy, always know their routes, and can be counted on to stay always sober. That's the safety promise of autonomous vehicles – not an absolute safety guarantee, but nonetheless a major improvement. From the Navlab project at CMU in the 1980's to "No hands across America" autonomous highway driving in the 1990's to present large-scale commercial endeavors, autonomous vehicles are the future of safe transportation.
- ***Workplace safety***: What if blue-collar workplace accidents could be predicted and thereby reduced? AI, in the form of historical workplace data analysis, correlating accidents with workplace conditions, safety inspections and behavioral indicators, can help do just that, as evidenced by joint work between our university and a company called Industrial Scientific, already helping improve worker safety in many companies across different industries such as construction, mining, manufacturing, etc.
- ***Massive multilinguality***: There are some 6,000 languages in the world, but only the top 2% or so by population and economic significance have translation software. What if there is a natural disaster requiring international assistance, or what if the US military needs to coordinate with local authorities for the other 98%? NSF, ARO and now DARPA are addressing this challenge to build information extraction from low-resource languages and rudimentary translation in a matter of days, rather than years.
- ***Game theory***: AI has proven itself at full-information games such as chess and go, and more recently at partial-information games such as Texas holdem poker, beating the human champions in each of these games. The same technology is used to optimize kidney exchanges, negotiations, and other partial knowledge high-payoff decision tasks.

The above examples represent a very small cross-section of a much larger AI application space. How about agriculture? AI is optimizing planting and fertilizing plans and schedules as well as creating robot-controlled tractors. Law enforcement? Voice and facial identification are helping forensic analysis. Music? AI-based accompaniment and AI-based tutors for different instruments. Healthcare? AI-based DNA analysis for risk factors, robo-surgery for super-human precision, analysis of patient records, and much more, are emerging. Finance? Hedge-fund analytics based on machine learning, investment risk, and so on. Education? AI based tutors in mathematics, language, and computer programming are emerging. There is nary an industry untouched by AI, and its impact will significantly grow and become more manifest in the coming years across the board.

Artificial General Intelligence (AGI): AI is rapidly embedding itself into just about every aspect of our lives; its utility is beyond question. However, how about the original AI dream of

creating general human intelligence? Many AI researchers have not given up on that goal, though many feel the path leads through the various components of intelligence before striving for a grand unification. Other AI researchers are content with narrow AI, that is, task-oriented AI and not general intelligence, since that is what drives our economic engine. Yet others, impatient to return to the AI-genesis goal of creating human-level intelligence coined the phrase "Artificial General Intelligence" seeking to leapfrog a component-based approach with a more holistic one.

Some AI researchers, such as this writer, view AGI as an aspirational goal, working to generalize AI methods from the narrow to the broad, from the specific to the general. For instance, consider transfer learning. If a person learns to drive a car, perhaps through instruction and considerable practice, and then she is asked to drive a van or a small truck, chances are she can do so, albeit with a few awkward moments adjusting to the height and size of the new vehicle. However if the software trained for a self-driving car were to be extracted and implanted onto a small truck, disaster may ensue – judging distances to other vehicles would be off due to the higher camera angle, the breaking distance and turning radius would not behave as expected, and even the lane-detection and following might fail. The difference is that humans transfer what they learned previously while compensating for the known differences, such as height, angles, size, vehicle responsiveness, etc. Transfer learning strives to do exactly that in an automated manner, learning what to keep, what to discard, what to modify and how to modify it across related tasks. Currently fielded AI approaches require retraining from scratch, insensitive to having learned a very similar task. Hence transfer learning is a small but important step towards AGI.

Another example is deep neural networks whose topology must be currently hand crafted by researchers for a given task, e.g. for face recognition vs reading MRIs, or for machine translation vs text mining. Each task requires determining how many nodes are required, arranged into how many layers, what type of layer-based combination functions, what kinds of connectivity among the layers, and so on. Researchers, including this writer, are striving to create self-configurable deep networks that change their structure automatically to optimize task performance. That is another small step towards AGI. Many other researchers are also investigating how to make AI more general, one step at a time, in many different ways and in different subfields such as in robotics, natural language processing, automated reasoning or machine learning.

However, a few researchers are less patient, trying to reach AGI by more direct means, and though not yet having achieved the desired breakthroughs, are nevertheless determined to pursue the proverbial gold at the end of the long rainbow.

Some common AI myths: Some popular beliefs or claims about AI deserve the label of "myth," as they are contrary to observations and informed opinion, including the following:

- ***One stop shop for all of AI!*** AI is not binary; it is not something one has or fails to have. The field continues to progress at a good pace; over time more sophisticated and capable

AI methods and systems are created. Any vendor claiming to be the ultimate provider of everything AI is selling a bill of goods or a pig in a poke. As more AI capabilities come online we must be ready to see whether and how to employ them, and we never know for certain where or by whom they will be created.

- ***AI = deep neural nets.*** Over its history many AI paradigms emerged, dominated the field, and then fell out of favor to newer methods: first-order logic, rule-based systems, the first coming on neural nets (1980s), symbolic machine learning, the second coming of neural nets (1990s), statistical machine learning, probabilistic reasoning, and now the third coming of (deep) neural nets. Actually, every AI paradigm leaves its mark, and many powerful systems use hybrid approaches, not just the latest method *du-jour*, even as convolutional and recurrent deep networks are proving to be very powerful.
- ***AI = suite of software tools.*** This is no more true than carpentry being just a box of carpenter's tools; one also needs the master carpenter. Surgery is not just a suite of scalpels, sutures and other surgical implements; one needs the experienced surgeon. AI tool suites are major enablers of novel AI applications, but the skilled AI practitioner is an integral part of the equation.
- ***AI is impossible.*** This used to be a common way for pundits to gain attention, including some philosophers, claiming that AI would never create a champion chess player, and would never understand human speech, and would never drive a vehicle in traffic. Over time all were proven wrong. Now the more common claim is that AGI is impossible. That one will be harder to disprove as AI systems become more general and more powerful, the proverbial goal posts will be moved from passing the Turing Test[1] to ever harder tasks as each is accomplished over time. Time and much research will tell whether true AGI is indeed achievable.
- ***AI will cause massive unemployment.*** A variant of this claim has been made after every major technological advance, but the opposite is typically true – there was more employment after the industrial revolution than before, or after the introduction of information technology and the internet. AI is already deployed in many industries, yet we have the lowest unemployment rate in recent times. Instead, AI will likely change the nature of work, displacing workers in some fields over time and creating other remunerative jobs. Rather than speculation, careful studies by economists, AI scientists and policy makers are required to make better predictions of labor market effects of AI – we simply do not know them in any detail. But it is fairly clear that the countries most advanced in AI technologies will reap its benefits, and others will be left behind, just like in earlier industrial revolutions, as the effects of AI will indeed be global.

[1] The most popular version of the test designed by Alan Turing, the famous code-breaking British mathematician, is whether a person asking questions of two hidden responders, one a human the other a machine, can tell which is which after a brief period of time based on their responses. If the questioner cannot tell them apart, then the machine is said to have passed the Turing Test.

Glimpsing into the future of AI: The most reliable prediction one can make relating to AI is extrapolating current trends into the near future:

- *AI will become increasingly ubiquitous in everyday life*, from general chatbots, to individualized health monitoring, to personal assistants. Turing award winner Professor Raj Reddy from CMU, predicts that personal assistants will evolve into "guardian angels", always-on AI system helping people prepare for meetings, monitoring news or social media for items of immediate interest, reminding users to exercise or to not touch that temping chocolate dessert. These "angels" or "cogs" as this writer prefers to call them will be driven by observing, learning, and receiving instructions from their users.
- *The power of underlying AI technologies will increase,* as machine learning, language technologies and robotic sensing improve due to more powerful GPUs, improved deep neural network architectures, better training algorithms that combine raw data with domain knowledge, and so on. These will increase the range of practical AI applications across the board.
- *AI will expand into new areas including cyber security,* where research is already establishing that zero-shot malware detection is possible via machine learning methods trained on prior cyberattacks and malware profiles. This is a major improvement to waiting for a new virus to infect many machines before viral signature is disseminated by the cyber protection provider. AI can also help in energy conservation, including smart-grid and smart-home power consumption, as well as optimizing renewable power installations via predictive local-climate models.
- *Artificial general intelligence will witness a resurgence,* not so much from a major paradigm switch in AI, though that is what its proponents desire, but rather from evolutionary forces in bringing to bear transfer learning, self-configurable neural networks, more flexible robotics, more natural and general dialog systems, and other improved AI methods coming together into larger systems of wider capabilities.
- *The demand for AI talent will far outstrip the supply.* Industries engage in bidding wars and employee poaching in attempts to fill open positions in AI, raising compensation levels well beyond any previously seen. Universities strive to increase the supply, but when industry attracts top university faculty – and this is happening with increasing frequency – it is very difficult to maintain, let alone expand, educational offerings.
- *AI in other countries will surge.* This is already happening in Russia, India, and especially China (see the following section), but the trend will expand to many more countries that realize AI is the crucial technology of tomorrow, and quickly ramp up efforts in research, education and especially industry.

- *AI ethics will become a central concern.* As with virtually all technologies, AI can be used for the benefit of human kind or its detriment. For instance, electricity powers our world, but also can electrocute; the internet is a major economic driver and wonderfully empowering by democratizing information but it is also used to spread hate-speech and enable cyber bullying. AI is no different, and we must consider ethical implications, where "we" refers to AI researcher and practitioners, ethicists, economists and policy makers working together. For instance we may want AI to help us establish and maintain healthy lifestyles, but we may not want that same AI to breach our privacy and inform others of our individual health problems.

US Dominance in AI: The United States has been the undisputed world leader in AI from its inception in 1956 through the early days of overoptimistic expectations that AI was just around the corner, to the realism that AI is difficult – much harder than rocket science – and its ultimate success in a wide variety of challenging tasks. In the 1980's Japan's Fifth Generation Project challenged the US AI dominance, and the 1990's Europe's Economic Commission attempted to surpass US AI efforts, but neither succeeded in large part due to the foresight and perseverance of the US public sector. Unlike the above limited-lifetime efforts, the National Science Foundation (NSF), the Defense Advanced Projects Agency (DARPA), the US intelligence community including the Intelligence Advanced Projects Activity (IARPA), and the research divisions of each of the armed services (ONR/NRL, AFRL, ARL/ARO) persevered in supporting AI through all of its growing pains, its successes, its more challenging times, and its arrival in the mainstream to become a top candidate for the most versatile and valuable industry of our century. The foresight to persevere, to strive to win the long game of AI, is a credit to the cited government institutions through multiple administrations and through remarkable bipartisan unity in the national interest. Hence, my top recommendation is kudos to these agencies; they do an excellent job, please extend them as much support as possible.

However, the US dominance in AI is being challenged like never before. Many countries are striving hard to improve their AI know-how, work-force, and industry, including China, Russia, Korea, Japan, Germany, the UK and India. Consider China, for instance, which has made AI a national priority. On May 4, 2018 CNBC reported "China is determined to steal the AI crown from the US and nothing, not even a trade war will stop it. China's 2030 plan envisions building a $1 trillion AI industry." Wired Magazine reports: "China will be the world's dominant player in artificial intelligence by 2030. This isn't a prediction by a researcher or academic, it's government policy from Beijing." Whereas these statements may be on more alarmist than reliable predictions, they clearly indicate Chinese intent. China's national priority is AI pre-eminence. Even General Secretary Xi Jinping is reported to have AI books on his shelf.

It is difficult to estimate the very substantial level of AI funding in China, but there some components include: 1) The city of Tianjin is committing $5 billion to support the new AI industrial park. 2) The Feb 20, 2018 statement in the Financial Times saying "*Last year almost half the global investment into AI startups went to China, up from a mere 11.3 per cent slice in*

2016", 3) On June 22, 2018 the South China Morning Post reported: "*China's Ministry of Science and Technology has funded at least eight AI-related research projects over the past six months to the tune of 2.73 billion yuan (US$430 million) from the central government budget*" and "*The China Academy of Sciences (CAS) which has over 300 labs and four national research centres, received over 2.7 billion yuan for its 11 fundamental science projects last year, although it's unclear how many of these are directly-AI related.*" China has already far surpassed the US in terms of patents granted for AI technologies, according to Quartz, May 2, 2018.

In contrast, the United States spends about $2 billion per year on AI research, according to OSTP, not counting substantial DoD spending on procurement that includes applied AI research. The bottom line is that both countries take AI very seriously and that China is striving and investing to achieve AI leadership. It will take a greater effort for our country to maintain international leadership in AI.

Some recommendations: The following suggestions address what this writer believes to be national priorities with respect to AI, whose neglect will be to our collective peril.

- ***Support the US government agencies which helped create and foster AI.*** NSF, DARPA, IARPA, parts of NRL/ONR, ARL/ARO, AFRL, etc. Without DARPA in the early days, AI would not exist in its present form. These agencies do a good job of funding and shepherding AI research. They deserve increased financial support to do their jobs even better, especially in light of the sudden large-scale international competition. No matter what other measures the US government takes with respect to AI, keep and grow the proven successful processes and agencies.
- ***Address the AI personnel shortage.*** We need to produce more AI researchers, especially more US-citizen or permanent-resident AI researchers. One way to do so is to provide scholarships for US students pursuing AI careers. One thought is that these could be funded by asking companies in an expanded merit-based H1B visa program to pay $30K or so per sponsored visa. Then if there are some 100K new H1B visas per year that would create a fund of $3 billion per year, which would provide something like 30 thousand scholarships. Whatever the mechanism or the numbers, we should be training more AI researches and engineers. The flipside is to retain more the foreign AI talent trained in US universities, which is also addressable by increasing H1B visas
- ***Address the exodus of AI faculty from universities.*** Industry is making offers that few AI faculty can refuse, with the inevitable loss of top university AI talent making it very difficult to teach and mentor new generations of students. We must stop eating our seed corn, though it is not clear how to change incentives to attract and preserve more top-talented AI faculty.
- ***Encourage AI at the undergraduate level.*** AI used to be for PhD students, then MS programs in AI or sub-disciplines of AI started including several at CMU. Now is the time to start AI as an undergraduate major, as we are experimentally doing at CMU.

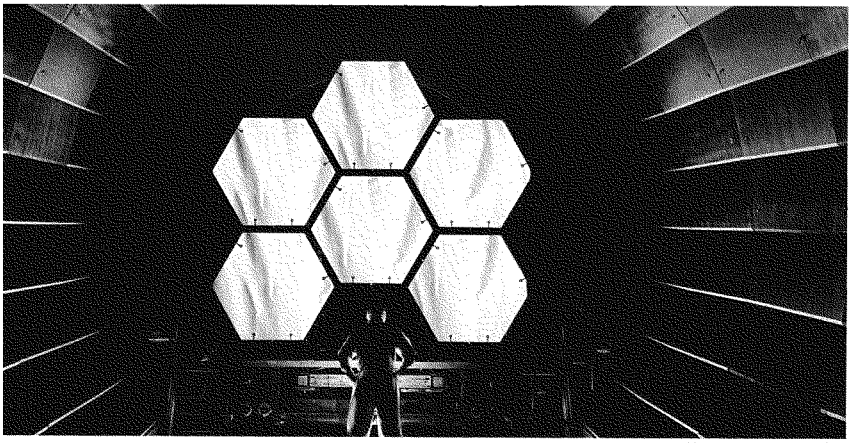
- ***Consider funding a national AI center.*** Other countries have national AI laboratories or are creating same, as a nexus between government, industry and academia. The US should do so as well, both as a vehicle for growing the volume and scope of AI activities but also to enable large-scale projects that require sustained collaboration among dozens of researchers, thereby balancing the smaller efforts supported by the NSF and other funding agencies. Such a center, though, should be stable for a long period of time to be truly productive, not a one-year flash-in-the-proverbial-pan.

Dr. Jaime Carbonell is the Director of the Language Technologies Institute and Allen Newell Professor of Computer Science at Carnegie Mellon University. He received SB degrees in Physics and Mathematics from MIT, and MS and PhD degrees in Computer Science from Yale University. His current research includes machine learning, artificial intelligence, deep neural networks, scalable data mining, natural language processing, and applications to finance, cybersecurity and computational proteomics. He invented proactive machine learning, including its underlying decision-theoretic framework, and has recently worked on new methods for transfer learning. He is also known for the maximal marginal relevance (MMR) principle in information retrieval, for derivational analogy in problem solving, for example-based machine translation and for machine learning in structural biology, and in protein interaction networks. Overall, he has published some 380 papers and books, and has supervised or is supervising some 74 PhD dissertations. He has received several awards for teaching and research, including the Okawa prize in 2015. Dr. Carbonell has served on multiple governmental advisory committees such as the Human Genome Committee of the National Institutes of Health, the Oakridge National Laboratories Scientific Advisory Board, the National Institute of Standards and Technology Interactive Systems Scientific Advisory Board, and the German National Artificial Intelligence (DFKI) Scientific Advisory Board. He has also co-founded three AI-based companies and serves on the board of Carnegie Speech.

REPORT SUBMITTED BY REPRESENTATIVE BILL FOSTER

New Democrat Coalition Economic Opportunity Agenda: A Future that Works

CLOSING THE SKILLS AND OPPORTUNITY GAPS



Skills gap
Refers to the supply-side mismatch between opportunities and skills, whereby employers are unable to fill positions because they cannot find qualified candidates.

The nature of work and the economy are changing. New jobs are being created, triggering a demand for technological skills across industries and all worker skill levels. Advancing technology is also disrupting and shrinking traditional middle-skill jobs in favor of new ones. These new middle- and high-skill jobs are clustering in and near certain urban areas,[1] leaving many Americans geographically disconnected from opportunity. Not only do the skills earned today have a shorter shelf-life than in the past,[2] but the traditional employer-employee relationship is changing as some work becomes more short-term and project based. As conditions continue to shift, American workers are finding it harder than ever to earn a good life.

Technology matters more than ever. As artificial intelligence (AI) becomes more integrated into the medical field, for example, radiologists are working alongside machines to improve patient care. Employers are unable to find enough qualified candidates to fill positions in emerging and growing industries, while middle-skill jobs made up 53 percent of the U.S. economy in 2015, only 43 percent of workers were trained at the level employers required.[3] Key drivers of this disparity are both the "skills gap" and the

1. Holzer, Harry (2015), "Job Market Polarization and U.S. Worker Skills: A Tale of Two Middles", Economic Studies, Brookings Institution
2. Eggers, William D., Hagel, John and Sanderson, Owen (2012), "Mind the (Skills) Gap", Harvard Business Review
3. Stephens, Rachael, "Mind the Gap: The State of Skills in the U.S.", Third Way

Opportunity gap
Refers to demand-side biases or hiring policies that place a greater emphasis on degrees over real world experience, causing qualified candidates to be overlooked.

Examples include degree inflation, *the exclusion of workers with criminal records, and exclusion of those with differing abilities or mental illness.*

"opportunity gap," which lead to worker shortages in many industries. This constrains economic growth, the earning potential and upward mobility of workers, and the creation of new industries and businesses.[4]

In many cases, workers are unaware of or unable to access opportunities to build new skills and become qualified for open positions in growing fields. However, employer biases for overqualified candidates, and a failure of the labor market to properly value all work, contribute to the persistence of these perceived shortages.

The New Democrat Coalition has identified several key challenges that are widening the skills and opportunity gaps:

- **Workers in certain industries and segments of the workforce are more vulnerable to technological disruption.**

  The labor market is becoming less mobile and dynamic. The workforce is aging, and new jobs are often concentrated in more expensive urban and suburban areas while rural and poorer communities are left behind.[5]

  Low-skill workers are significantly more vulnerable than high-skill workers to disruption and are less likely to have access to tools to help them find new work.

- **Firm-worker matching needs to be improved.**

  Many employers report that candidates do not have the skills necessary for the jobs available in part because:

  - Hiring requirements do not always match the skills necessary to perform a job. This often manifests in businesses requiring higher educational attainment for candidates than necessary.
  - Many skills learned outside of traditional pathways are difficult to effectively communicate on a resume.
  - Some workers do not possess the necessary skills to perform jobs in growing industries and fields.
  - Employers are investing less in job training for their workers than they did in the past.[6]
  - Many Americans, including those with criminal convictions or fighting addictions, are automatically or effectively disqualified from many jobs. Because the criminal justice system incarcerates a disproportionate number of people of color, this is a pronounced challenge in minority communities.

4. Stephens, Rachael, "Mind the Gap: The State of Skills in the U.S.", Third Way
5. Bloomberg (2017), "New America and Bloomberg Announce Commission's Findings on the Future of Work", Press Release
6. Executive Office of The President (2015), "Economic Report of the President" with the "Annual Report of the Council of Economic Advisors", Transmitted to Congress (Obama Administration)

Career and Technical Education (CTE)
A term applied to schools, institutions, and educational programs that specialize in skilled trades, applied sciences, modern technologies, and career preparation.

CTE aims to provide students with in-demand technical skills, as well as to prepare them for post-secondary degrees and credentials in in-demand careers.

Examples of CTE programs include technical secondary and post-secondary classes, employability classes, and on-the-job training through internships and apprenticeships.

- **Existing models of education are not meeting the needs of some students and workers.**

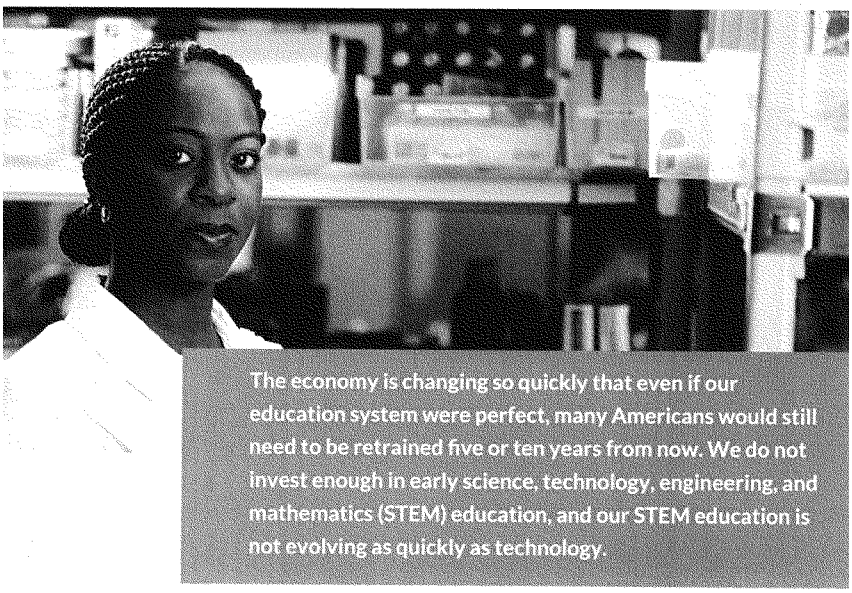
  Some students and workers are not learning the skills they need to compete in the changing economy:

  - The economy is changing so quickly that even if our education system were perfect, many Americans would still need to be retrained five or ten years from now. We do not invest enough in early science, technology, engineering, and mathematics (STEM) education, and our STEM education is not evolving as quickly as technology.
  - We must also recognize the importance of, and continue to invest in, education that develops "soft" skills, such as critical thinking, creativity, and interpersonal skills that cannot be easily automated[7] and are transferable across all industries and job functions.
  - We do not invest enough in training workers likely to be displaced by automation for work in expanding industries.
  - Various economic drivers have caused higher education costs to continue rising faster than inflation and wages.
  - Current professional licensing and credentialing practices are often burdensome and restrictive.
  - Career and technical education (CTE) can provide pathways to middle-skill, middle-income careers, but not enough students pursue this secondary and post-secondary educational option.
  - Our current culture of education does not place enough value on technical education or lifelong learning.
  - Parents, academic advisors, and students do not possess the necessary information to understand the value of apprenticeships, community college, and CTE.
  - Existing CTE programs do not always prepare students for the workforce or align with local employers' needs.

To better understand the challenges and opportunities of 21st century skills training and workforce development, the New Democrat Coalition's Future of Work Task Force held a public forum[8] on July 19, 2017. It featured Harry Holzer, Professor of Public Policy at Georgetown; Byron Auguste, CEO and Co-founder of Opportunity@Work; Portia Wu, Director for Workforce Policy for Microsoft; and Julie Peller, Executive Director of Higher Learning Advocates. The panelists discussed the growing skills

7. Nedelkoska, L. and G. Quintini (2018), "Automation, skills use and training", OECD Social, Employment and Migration Working Papers, No. 202, OECD Publishing, Paris
8. New Democrat Coalition (2017), "How Congress Can Help Americans Succeed in the Economy of the Future", Medium

gap and how to close it. Task Force Members and staff have also met with policy, industry, and thought leaders in the field, including Colorado Governor John Hickenlooper; Rhode Island Director of Labor and Training Scott Jensen; Swiss Ambassador Martin Dahinden; Director of The Hamilton Project, Jay Shambaugh; and Professor of Economic Policy at the Harvard Kennedy School, Jason Furman. From these discussions, we have identified several key principles and policies to give all Americans the opportunities to succeed in the changing economy.



The economy is changing so quickly that even if our education system were perfect, many Americans would still need to be retrained five or ten years from now. We do not invest enough in early science, technology, engineering, and mathematics (STEM) education, and our STEM education is not evolving as quickly as technology.

PRINCIPLES FOR EXPANDING OPPORTUNITIES IN THE CHANGING ECONOMY

The New Democrat Coalition believes Americans are looking for the opportunity to earn a good life. To better prepare all Americans for the changing economy, we must change our approach to education. Learning must become lifelong, and our educational systems updated with input from the public, private, and non-profit sectors. We must either adapt and take advantage of this moment, or risk being left behind.

Investing in human capital and the skills of our workers is paramount to driving wage growth, productivity, and improving the standard of living for all Americans.

Investing in human capital and the skills of our workers is paramount to driving wage growth, productivity, and improving the standard of living for all Americans. The New Democrat Coalition recommends several policies to provide American workers with more opportunities to succeed organized around three core principles:

Give every American access to the tools to succeed in the 21st century economy

Advancements in technology have already caused consequential shifts in the kinds of jobs available to workers. These shifts will bring opportunities for some, but could leave others behind. Educational attainment is key to success in the new middle-skill jobs in expanding industries such as healthcare, information technology (IT), construction, and high-tech manufacturing.[9]

Workers will not always be able to afford the investments needed to keep pace with changing skills demands, and many companies do not have the capacity or incentive to invest in training workers who may take their skills elsewhere. Policymakers can bridge the gap by encouraging the private sector to make further investments in their workforce.

We must expand high-quality, shorter-term, flexible educational opportunities that offer transferable, widely applicable skills and credits and clear career pathways.

Remove obstacles for Americans to access middle-skill jobs and close the opportunity gap

Important to bridging the skills gap is better identifying the talent that exists within the current workforce. In some instances, employers screen out qualified candidates because they are unable to properly evaluate their experience or establish screens that unwisely preclude an entire category of workers.

Yet some careers will require workers to pursue additional training. As career and technical training, credentialing, and certification testing programs are developed and updated, workers should be able to transfer earned skills between states, industries, companies, and educational programs.

Create a culture of lifelong learning and make it easier for Americans to invest in their future

Our higher education system was designed for full-time students, but today's student looks very different. Between 1989 and 2008, 70 to 80 percent of college students were working while taking classes, with 40 percent of undergraduates and 76 percent of graduate students working full-time.[10] Research has found that 26 percent of under-

9. Stephens, Rachael, "Mind the Gap: The State of Skills in the U.S.", Third Way
10. Carnevale, Anthony, Smith, Nicole, Melton, Michelle and Price, Eric W. (2015), Georgetown University McCourt School of Public Policy, Center on Education and the Workforce "Learning While Earning: The New Normal"

Upskilling
Refers to the teaching of or learning of additional skills by workers. This can be done for personal and professional development, as well as to adapt to changing skills demands within a profession. The Aspen Institute has partnered with Upskill America to encourage employers to invest in upskilling for their employees.

graduate students are raising children.[11] Over a quarter of students enrolled at a post-secondary level take distance learning classes, such as online courses.[12] Furthermore, as technological progress accelerates, the shelf-life of applicable workforce skills, or the period of time skills remain relevant and applicable, has shrunk to five years and will likely continue to shrink.[13] The skills a worker learns before entering the workforce often do not apply as technology and shifts in the labor market change the nature of their work.

When a student—either before entering the workforce or far into their career—does pursue a college degree, they should have options that can conform to their schedule.

Education should be lifelong, stackable, and exist on a continuum.

Education should be lifelong, stackable, and exist on a continuum. In other words, the skills and knowledge a worker receives should be universally recognized and interoperable with additional education and training a worker may receive in the future. For example, a student or worker who participates in career training should have the opportunity to apply those credits to other training programs, or pursue an Associates or Bachelor's degree. Similarly, a worker or student with a two- or four-year degree should have the flexibility to acquire new skills in credentialing and training programs as needed. As the shelf-life of technical skills continues to decline, new technology is integrated into workstreams, and new skills become in demand, companies and workers should adapt by developing and participating in frequent, short-term upskilling opportunities.

11. Gault, Barbara, Reichlin Cruse, Lindsey, Noll, Elizabeth, Sykes, Mary, Talbourdet, Lauren and Eckerson, Eleanor (2016), "Child Care for Parents in College: A State-by-State Assessment", Institute for Women's Policy Research
12. National Center for Education Statistics (2016), "Fast Facts: Distance Learning (Fall 2014)", U.S. Department of Education
13. Eggers, William D., Hagel, John and Sanderson, Owen (2012), "Mind the (Skills) Gap", Harvard Business Review

POLICY ACTION PLAN TO GIVE AMERICANS THE TOOLS TO COMPETE IN THE CHANGING ECONOMY

Integral to achieving these goals will be expanding access to and availability of high-quality, affordable, flexible, skills-oriented learning and education, such as apprenticeships and credential and certification training and testing. The lifelong student should be the focus, allowing new and existing workers to build on a foundation of skills that adapts to labor market changes. Today's workers may have to upskill as frequently as every five years.

It must become the norm for employers and workers to approach learning as a lifelong experience and a priority investment.

It must become the norm for employers and workers to approach learning as a lifelong experience and a priority investment. Skills received must be transferable, flexible, affordable, and provide quantifiable value to students and workers. Employers also must

Soft skills
Sometimes referred to as noncognitive skills or employability, are a combination of social, emotional, and behavioral skills such as communication, teamwork, perseverance, problem solving, emotional judgement, and ethics

Hard skills
Sometimes referred to as cognitive skills, are teachable abilities that are quantifiable, including technical skills, math skills, and reading skills.

update their hiring practices to recognize the existing skills in the workforce. Sometimes, a requirement on a resume outweighs an evaluation of a candidate's ability to do a job. For the labor market to perform properly, and for the good of the economy, employers, and American families, we must value the skills and experience of workers, not just academic pedigree. We will offer several policy proposals to help build a culture of lifelong learning that expands the options for students and workers without committing them to a rigid career path. The nature of learning must change with the nature of work.

Give every American access to the tools for success in the 21st century economy:

Provide students and workers with universal basic skills: To prepare our workers for the future, we must provide all Americans with a set of universal basic skills. In the 21st century economy, workers should be prepared with critical thinking, digital, and job-specific skills – including business fundamentals and industry knowledge, as well as personal skills.[14]

- **Invest in STEM education.** Employers value the critical thinking skills that workers obtain through liberal arts educations, but we must strengthen our STEM pipeline. Technical skills, such as computer science, are now a fundamental requirement across many fields. Giving students early experience with computer science and coding could help prepare them with fundamental skills for a variety of careers.

  A report from Third Way found that a diverse range of industries face qualified-worker shortages, particularly in healthcare and social assistance (e.g. nursing), professional and business services (e.g. tech services), and educational positions in STEM.[15]
- **Emphasize and value "soft" skills, which are harder to automate.** We need to focus on teaching students "soft" skills that cannot be automated, including people skills, creativity, foreign language skills, and critical thinking skills.[16] Workers who possess both "soft" and "hard" skills are particularly competitive.[17] And while investment in STEM is important, STEM fields are not immune from automation. With our aging popu-

14. Stephens, Rachael, "Automate This: Building the Perfect 21st-Century Worker", Third Way Report
15. Stephens, Rachael, "Mind the Gap: The State of Skills in the U.S.", Third Way
16. Nedelkoska, L. and G. Quintini (2018), "Automation, skills use and training", OECD Social, Employment and Migration Working Papers, No. 202, OECD Publishing, Paris
17. Whitemore Schanzenbach, Diane; Nunn, Ryan; Bauer, Lauren; Munford, Megan; and Breitwieser, Audrey (2016), "Seven Facts on Noncognitive Skills from Education to the Labor Market", The Hamilton Project, Brookings Institution

Experiential learning
Refers to education that develops skills and knowledge through direct experiences outside of a traditional academic setting. This includes apprenticeship programs, internships, or other hands-on learning.

Career matching
Involves providing better data and transparency to help job seekers identify in-demand skills for different fields, as well as helping employers access and identify qualified candidates.

lation and more women joining the workforce, there will be a dramatic increase in the demand for home healthcare workers, child care workers, and other types of home care work—fields which often require both kinds of skills.

- **Provide work-based learning during high school** to give students more exposure the labor market through experiential learning and better career matching. Colorado, for example, established Launch My Career Colorado,[18] a tool to help students and workers assess the value of different degrees and certificate programs to achieve their career and education goals. The tool can be used to identify growing job markets and the degrees and certification programs that will appropriately prepare students. It can also be used to compare future earnings against the investment required to obtain the necessary degrees.

Academic advisors, parents, and students must have access to information about their options and the value of apprenticeships, community college, career training, and other educational opportunities.

Invest in career training programs at the federal, state, and local levels: Career and technical education, including credentialing programs and certification testing at community colleges and through innovative, newer programs, present excellent opportunities for students and workers of all ages and skillsets to learn and grow. We must directly invest federal dollars in the expansion of programs that have successfully assisted students and workers and encourage state and local investments tailored to local needs.

- **Modernize federal investment in Career and Technical Education (CTE):** In June 2017, the House unanimously passed Rep. Raja Krishnamoorthi's bipartisan Strengthening Career and Technical Education for the 21st Century Act, which will help more Americans get relevant skills by making upgrades and investments in secondary and post-secondary career education. This includes career-training programs, which can help more Americans gain the knowledge and skills they need. It also aims to improve alignment between educational organizations and employers, and encourages a focus on employability skills, work-based learning opportunities, and meaningful credentialing. The Senate must pass this important legislation.
- **Reorient post-secondary education and workforce development programs to prepare workers with career-ready skills:** Programs focused on skills training, post-secondary education, and workforce development were created for an economy in which people often worked their entire careers at one company or industry.

18. Launch My Career Colorado

Post-secondary education needs to be adapted to the careers of the 21st century by working with the private sector to develop career-ready curricula and programs that are updated continuously based on the success of graduates and changing employer needs

Post-secondary education needs to be adapted to the careers of the 21st century by working with the private sector to develop career-ready curricula and programs that are updated continuously based on the success of graduates and changing employer needs.

Evidence shows that the most successful workforce development programs teach skills in high demand for local markets.[19] Partnerships between post-secondary educational institutions, such as community colleges, and industry can increase access to information and data on the skills needed for in-demand fields, and can give training programs the resources and infrastructure to adapt to changing labor market trends.

- **Invest in apprenticeships developed through partnerships between industry and educational institutions.** Skills acquired through registered and more innovative apprenticeships must be portable. To achieve this, employers and educational and training institutions must be engaged in developing standards for these programs recognized across industry and other educational and training institutions. Working closely with the private sector and labor unions, community colleges can play a valuable role in facilitating apprenticeships, ensuring quality and establishing an interoperable system of credentials. We believe all apprenticeships should culminate in a degree, nationally recognized or industry-wide certification or credential, and that federal investment should expand models of success.

Several New Democrat Coalition Members have introduced legislation investing in post-secondary education and to encouraging employers and state and local governments to invest in apprenticeships and workforce development, including Rep. Seth Moulton's CHANCE in Tech Act, Rep. Terri Sewell's Workforce Development Tax Credit Act, Rep. Annie

19. Guvenen, Faith (2018), "Stagnation in Lifetime Incomes: An Overview of Trends and Potential Causes", The Hamilton Project, Brookings Institution

Kuster's Workforce Development Investment Act, Rep. Rick Larsen's Youth Access to American Jobs Act, and Rep. Norma Torres's Job Opportunities Between our Shores (JOBS) Act.

Community colleges play a critical role in preparing workers for many middle-skill jobs.

Increase federal funding to community colleges in exchange for greater accountability: Community colleges play a critical role in preparing workers for many middle-skill jobs. Increased funding to community colleges delivering career-ready skills should be tied to strengthening their occupational and workforce programming. Additional funding should be based on programs oriented toward student success, including programs that are targeted to high-demand fields and those that have support services to help disadvantaged and underperforming students. Federal resources should also take into account measures of success including job-matching rates for students after graduation. This will require accountability measures focused on outcomes, such as completion, post-graduation employment, earnings, and the skills and knowledge of students.

Unfortunately, because community college students often face significant financial and personal challenges, they often are unable to complete their degrees.[20] While completion rates for students in bachelor's degree programs grew from 55.4 to 59.6 percent from 2002 to 2014, they actually declined for two-year institutions from 30.5 to 27.9 percent[21]. The Department of Education can play a role in helping community colleges build "guided pathways," or focused student assistance to encourage course completion and career pathways, to help students complete their degrees or transfer into more appropriate programs.

Remove obstacles to access middle-skill jobs and close the opportunity gap:

Expand competency-based hiring: "If you can do the job, prove it and we will hire you." What if this were the motto for more middle-skill jobs? For many employers, the history listed on a worker's resume holds greater significance than proof they can do the job. Organizations like Opportunity@Work engage with employers across the country to reform hiring practices and help workers access affordable pathways to gain the skills they need.[22]

20. Deming, David (2018), "To increase college completion rates, invest directly in public postsecondary institutions", Brookings Institution
21. Deming, David (2017), "Increasing College Completion with a Federal Higher Education Matching Grant", The Hamilton Project, Brookings Institution
22. Opportunity at Work

Degree inflation
Refers to employer preference or requirements for four-year degrees for jobs traditionally held by middle-skill workers. Harvard Business School released a report that further investigates this trend.

We must lower unnecessary barriers for workers and enable them to access jobs based on demonstrated competency.

We must lower unnecessary barriers for workers and enable them to access jobs based on demonstrated competency. Expanding this approach to hiring will help give entry-level and low-skilled workers access to middle-skill jobs. According to a report by the analytic software company Burning Glass, 65 percent of open positions for executive secretaries and assistants require a bachelor's degree even though only 19 percent of workers currently in those positions hold such a degree. This degree inflation[23] occurs in other fields like IT help desks or construction supervisor positions. It is bad for workers and bad for employers who ultimately take longer to fill positions with inflated requirements.[24] This trend may begin to subside as the labor market tightens, forcing employers to reevaluate hiring criteria. But it is clear that a bias for higher educational attainment does play a real role in preventing workers from accessing jobs for which they are qualified.

In this area, the federal government can lead by example. The Office of Personnel Management (OPM) should ask federal agencies to review classifications of all positions and, where possible, scale down unnecessary education requirements with competency evaluations to measure necessary skills.

Congress can also authorize competitive grant programs in workforce and economic development to reward states that reduce their four-year degree requirements and other restrictive hiring requirements, supplementing them with skills-based hiring practices. These hiring practices give Americans with less academic pedigree a fair shot to be hired into government and government contractor career paths.

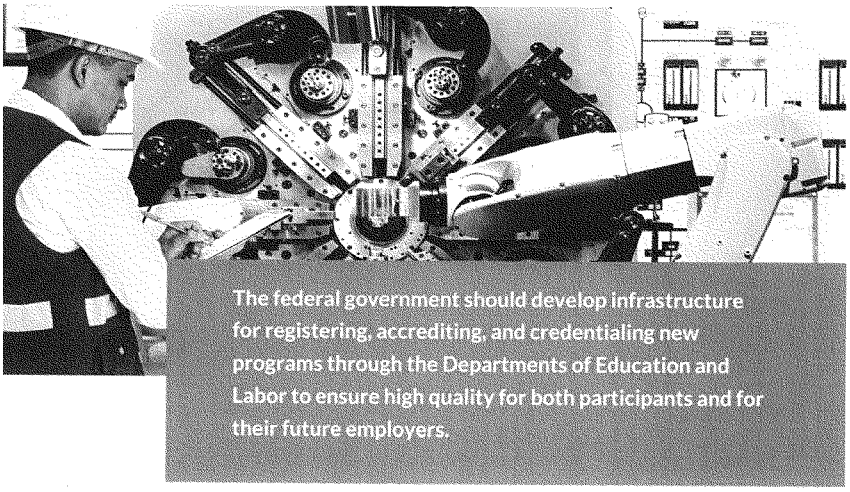
Accelerate the deployment of nationally-recognized credentialing standards: Credentialing standards and certification testing programs must be developed with private, public, and educational institution stakeholder involvement. Support from the institutions that train workers and the industry groups and companies that employ them will enable credentialing requirements to more quickly adapt to evolving technology and market demands, and will ensure credentials are widely

23. Deutsch, T., Blackwood, D., Eshun, T., & Alexanian, Z (2017), "New Report: Degree Inflation Hurting Bottom Line of U.S. Firms, Closing Off Economic Opportunity for Millions of Americans", Press Release, Harvard Business School
24. Burning Glass Technologies (2014), "Moving the Goalposts: How Demand for a Bachelor's Degree Is Reshaping the Workforce"

recognized. Nationally and industry recognized credentials, certifications, and licensing will promote mobility by enabling workers to more easily move between companies, industries, and across state lines while broadening the pool of qualified candidates.

The federal government should develop infrastructure for registering, accrediting, and credentialing new programs through the Departments of Education and Labor to ensure high quality for both participants and for their future employers. Other strong models include standard credentials developed in industry associations and consortiums, and state and military certifications and programs.

Veterans trained and equipped with skills for high-demand jobs face the challenge of entering the civilian workforce where their training is often not recognized under state and federal certification standards. In the 114th Congress, Rep. Denny Heck introduced the Collaborating for Economic Reintegration of Troops (CERT) Act, which would establish a commission to develop new, uniform licensing and certification standards to help veterans transition to civilian life and high-demand occupations, while avoiding conflicting state-based occupational licensure and experience requirements. A veteran who has driven a military truck under fire halfway across the world can probably handle driving a truck at home.



Ban the box

Refers to policies that prevent applications for employment, housing, benefits, insurance, loans, and other services from asking if an applicant has ever been convicted by a court. The goal of these policies is to highlight an applicant's qualifications without the stigma of their arrest record.

Expand the pool of eligible workers: Many Americans are automatically disqualified from or overlooked for many job opportunities. Millions of Americans are prevented from accessing good jobs due to past criminal convictions or arrest records. More must also be done to include women and people of color who are underrepresented in STEM fields. In addition, workers in many states may lose their professional licenses and ability to work if they fall behind on their student loans. We must implement policies that give more Americans more opportunities, not less.

- **Provide a fresh start and value potential.** According to the National Employment Law Project, there are approximately seventy million Americans with prior arrest records or criminal convictions.[25] Many of these individuals are frequently rejected by employers despite their qualifications. This is particularly troubling in cases in which crimes committed were minor, occurred a long time ago or have been remediated through the criminal justice system.

  Studies show that men who have served time in prison or have been convicted of a felony are more likely to be unemployed, regardless of age, ethnicity, or educational attainment. While research estimates that eight percent of the U.S. population have a felony conviction,[26] the Department of Justice does not maintain a national database of this population, making it incredibly difficult to understand the economic impacts of a felony conviction. Rep. Seth Moulton petitioned the Department of Justice and the Census Bureau to aggregate data to better understand the institutional barriers that exist for this population.

  Thirty-one states and hundreds of cities and towns across the country have adopted "ban the box" policies so that employers consider a job candidate's qualifications first, without the stigma of an arrest or criminal record. However, in many cases this has resulted in decreased employment for black and Hispanic men without criminal records.[27] Congress must pass comprehensive criminal justice reform to ensure those seeking to better themselves and provide for their families can to do so through meaningful employment.

- **Protect worker's ability to earn.** In twenty-two states, those who default on their student loans have their professional and/or driver's licenses suspended. These individuals often rely on their licenses to either perform their work or commute to jobs, which provide income to pay back their

25. National Employment Law Project, Ensuring People With Convictions Have a Fair Chance to Work Campaign
26. Flurry, Alan (2017), "Study estimates U.S. population with felony convictions", UGA Today, University of Georgia
27. Doleac, Jennifer (2016), "'Ban the Box' does more harm than good", Op-Ed, Brookings Institution

loans.[28] Congress should act, potentially by amending the Higher Education Act as it is reauthorized this Congress, to prohibit this practice and ensure that workers are still able to earn, rather than forcing them further into debt.

- **Enable access to work.** More must be done to attract and retain diverse talent in traditionally difficult to access careers. Because of historical biases and other systemic factors, women and people of color lack representation in STEM fields.

Research has shown that improving diversity not only expands opportunity for workers but also improves business performance.

Research has shown that improving diversity not only expands opportunity for workers but also improves business performance. Furthermore, many workers—such as those who are primary caregivers or have disabilities—are prevented from participating in the labor force, widening economic inequality and the gender wage gap.

- **Provide access to job training for those most vulnerable.** Currently under the Supplemental Nutrition Assistance Program (SNAP), adults without dependents can receive no more than three months of nutrition assistance within three years if they are not in an employment training program or working at least 20 hours a week. But states are not required to offer their unemployed citizens an opportunity to participate in a job training program. The SNAP Work Opportunity Act, introduced by Rep. Suzan DelBene, would apply the three-month limit for eligibility only if a person who could not find a job was offered a training or assistance opportunity. This ensures those who are looking for work will no longer be penalized because their state does not offer job training or a work program.

Create a culture of lifelong learning and make it easier for workers to invest in their future

Reform federal worker training and dislocation programs: The federal government spends billions of dollars on programs to help unemployed workers find new

28. Silver-Greenberg, Jessica, Cowley, Stacey and Kitroeff, Natalie (2017), "When Unpaid Student Loan Bills Mean You Can No Longer Work", New York Times

work. In 2010, the U.S. Government Accountability Office (GAO) found that there were 47 employment and training programs with combined appropriations of $33.8 billion.[29] One of the main federal programs that helps displaced workers is Trade Adjustment Assistance (TAA), which was established to retrain workers displaced by globalization. However, many reports have called into question the effectiveness of programs like TAA. Indeed, the current system does not cover contingent workers, can discourage retraining, and can discourage dislocated workers from starting new businesses.

Workforce assistance programs, such as TAA, should be expanded to address displacement due to broader economic trends, such as technology. Programs should not necessarily distinguish between the causes of displacement, but rather the impacted individuals and communities.

Federal workforce programs should be flexible, innovative, and responsive to changing workforce needs, allowing workers to transfer skills across companies and industries.

Federal workforce programs should be flexible, innovative, and responsive to changing workforce needs, allowing workers to transfer skills across companies and industries.

The government should audit existing workforce development programs, working with state and local governments and industry, to identify and expand the most effective programs and delivery models. Less effective programs should be reformed or replaced by new approaches that support job-seekers more effectively and that can attract employer investment.

Make upskilling and going back to school easier: In today's economy, workers need to continually upskill in order to succeed. To achieve this, flexibility is key: certification testing and skills training programs should be built around an accessible delivery model with clear, strong career pathways, adaptable to the schedules of working individuals and students. In order to upskill, many workers need additional support such as child and elder care.

Because many workers do not live near educational opportunities, we also need

29. U.S. Government Accountability Office (2011, January), "Multiple Employment and Training Programs."

to expand access to broadband to assist individuals in developing skills and earning credentials. Furthermore, for students pursuing two- and four-year degrees, we must incentivize educational institutions to develop more programs with schedules and formats that work for working students.

Incentivize employers to invest in worker upskilling: Employers used to invest more in training their workers, but over time, that has declined precipitously.[30] There has been a recent trend among some employers reinvesting in worker training, and public policy can help speed-up these investments through incentives. For employer based training to provide the most value to workers and employers, they should be developed in collaboration with industry stakeholders and educational institutions. Credentialing and certification programs built in partnership with industry can be more responsive to labor market demands without sacrificing the value of student-oriented education. Training programs developed through sector-based partnerships between industry, community colleges, and other organizations have proven particularly successful.[31]

We should develop policies that reward companies for good behavior: valuing human capital and investing in their workers.

We should develop policies that reward companies for good behavior: valuing human capital and investing in their workers

- **The tax code** should be realigned to better incentivize companies to hire and invest in people rather than favoring capital expenditures. We recommend:

    i. Lifting the current cap on employer contributions to outside employee educational and training costs. Currently, employers can deduct up to $5,250 in training costs at universities, technical schools, and community colleges per employee each year.

30. Executive Office of The President (2015), "Economic Report of the President" with the "Annual Report of the Council of Economic Advisors", Transmitted to Congress (Obama Administration)
31. Holzer, Harry (2015), "Higher Education and Workforce Policy: Creating More Skilled Workers (and Jobs for Them to Fill)", Economic Studies, Brookings Institution

ii. Allowing employees who do not itemize their deductions to deduct training and certification costs from taxable income.

iii. Reforming the tax code to encourage employers to provide student loan repayment for employees by allowing student loan payments to be fully tax deductible. This Congress, Rep. Suzan DelBene helped introduce the Higher Education Loan Payments for Students and Parents Act, or HELP for Students and Parents Act, which would allow employers to contribute to their employees' student loan debt or college savings accounts pretax, and provide a tax credit to employers based on 50 percent of these contributions. Rep. Scott Peters introduced the Employer Participation in Student Loan Assistance Act in 2018 to incentivize employers to help their workers' pay down student debt through a tax-exempt employer benefit.

iv. Senator Mark Warner and Rep. Raja Krishnamoorthi were joined by several colleagues in introducing the bicameral Investing in American Workers Act this Congress, which encourages employers to invest more in quality skills training for their workers. The legislation would create a tax credit for increased training expenses directed at lower- and moderate-income workers.

- **Training programs:** In Rhode Island, the Department for Labor and Training (DLT) has partnered with companies to directly solve the labor pipeline challenge through the Real Jobs Rhode Island program.[32] Companies hire candidates lacking the necessary skills or credentials, but who possess other valuable qualities. The DLT then either directly trains or pays for the training of these hires using a combination of foundation, state, and federal money. Other states have established similar programs. We encourage states to build training programs that partner with industry to identify and fill labor market needs.

We should facilitate the creation of worker-owned, employer-matched savings plans called Lifelong Learning Accounts.

Establish lifelong learning savings accounts to help workers save money for education and job training. We should facilitate the creation of worker-owned, employer-matched savings plans called Lifelong Learning Accounts. The federal

32. Rhode Island DLT Real Jobs Rhode Island program

government could model policy based on the lifelong learning accounts of Maine and Washington, which grow over time through payroll deductions and encourage employer contributions through outcome based tax incentives. Rep. Derek Kilmer will reintroduce the Skills Investment Act to establish a national lifelong learning account model.

Support and encourage new ways to finance education, especially as Congress works to reauthorize the Higher Education Act and considers how to make post-secondary education more affordable, accessible and accountable.

The cost of higher education has increased 63% since 2006.

- **Reform Pell Grants and expand the use of federal financial aid.** The cost of higher education has increased 63 percent since 2006, outpacing available federal financial aid and making college unaffordable for many Americans.[33] In general, more federal financial aid, whether through Pell Grants or other programs, should be made available to keep pace with these rising costs associated with education. Congress should allow Pell Grants to be used year-round and to pay for additional options, like high-quality CTE, credential programs, and other innovative post-secondary education options. Only programs that are providing value to students and taxpayers and have proven they are meeting labor market demands should be eligible for federal financial aid. To ensure clear accountability measures are in place, and to prevent the use of Pell Grants by for-profit programs that offer little or no value to students, high-performing programs should be registered and accredited with the Department of Labor and the Department of Education.
- **Expand flexibility of 529 College Savings Accounts to include more post-secondary options.** This Congress, Rep. Donald Norcross introduced the Opening Paths to Invest in Our Nation's Students (529 OPTIONS) Act, which would allow students and workers to use a 529 account for costs associated with apprenticeships and associated credentials and certifications. This approach can be expanded for non-apprentices enrolled in certification and credentialing programs.
- **Bring more innovation into financing and lessen the burden on students.** A finance mechanism to consider is the Income Share Agreement (ISA) model, in which investors or funds finance a predetermined amount of a student's education costs in exchange for a set percentage of their future income over a specified period of time. Structured properly, ISAs can mitigate the financial downside risk for students by reducing reliance on debt and providing a grace period and possible repayment forgiveness

33. Bureau of Labor Statistics, The Economics Daily, "College tuition and fees increase 63 percent since January 2006 on the Internet", U.S. Department of Labor

if a student's income is below a certain level. ISAs can also increase the accountability of post-secondary educational institutions by linking repayment to student outcomes. No student should go bankrupt pursuing their education. The federal government could appoint a regulatory body to oversee ISAs and ensure that consumer protections and quality assurances are in place.

Purdue University's Back a Boiler program[34] is an ISA model that supplements traditional financial aid to students from a blind fund within a larger trust developed by investments and donations from alumni and other investors. It is critical that more work be done assessing and implementing appropriate results-oriented, risk-sharing educational financing models. Reps. Jared Polis, Seth Moulton, Terri Sewell, and Kyrsten Sinema have cosponsored the ISA Act supporting the expansion of ISAs and establishing more federal oversight.

Make higher education more effective and accountable: We need new policies that focus on higher education accountability, transparency, and academics while also streamlining cumbersome and costly regulations and cutting red tape. For example, Rep. Jared Polis's College Transparency Act, introduced this Congress, would lift the ban on the sharing of student-level data at educational institutions, allowing students and families to access information that would help them to select the right school.

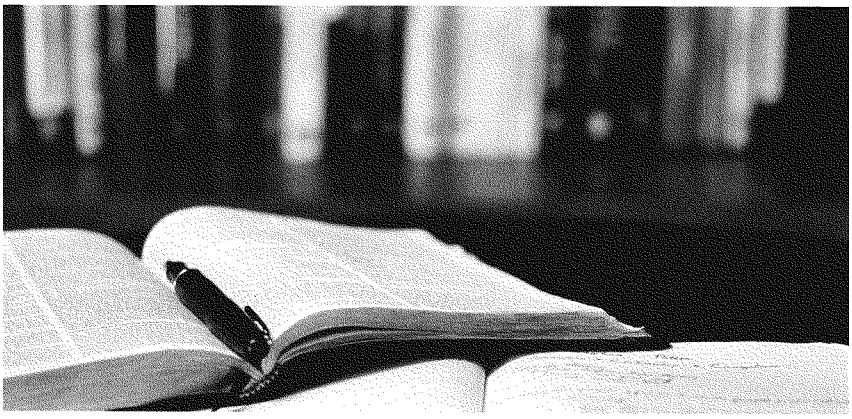
34. Purdue University, Back a Boiler

Federal policy should also leverage the power of private and public data to help educational institutions and private-public partnerships shape training and higher education curricula, and provide more information to workers, students, and their families about outcomes after graduation. In the future, data should be able to answer: What career paths and programs helped students and workers best adapt to automation and artificial intelligence? What additional skills did workers most need, and how did they acquire them?

Congress is currently working to reauthorize the Higher Education Act. This presents an opportunity to ensure our nation's institutions of higher education are providing value by equipping students with the knowledge and skills they need to be successful.

The Coalition will continue to explore innovative policies that support the development and education of our students and workers as they prepare for the changing economy. Addressing the skills and opportunities gaps will require collaboration between policy makers, industry leaders, and educational institutions.

Our approach must be holistic, encouraging companies to reform hiring practices, giving workers the skills they need, making education and training accessible and affordable, and raising the prestige and cultural value of educational pathways that focus on career ready skills. Everyone deserves a chance at the American Dream.



REFERENCES

Bloomberg (2017, May 16). "New America and Bloomberg Announce Commission's Findings on the Future of Work," Shift: The Commission on Work, Workers, and Technology.

Bureau of Labor Statistics, U.S. Department of Labor (2016, August 30). The Economics Daily, "College tuition and fees increase 63 percent since January 2006 on the Internet."

Burning Glass Technologies (2014, September) "Moving the Goalposts: How Demand for a Bachelor's Degree is Reshaping the Workforce."

Carnevale, A., Smith, N., Melton, M., & Price, E. (2015). "Learning While Earning: The New Normal," Georgetown University. Center on Education and the Workforce.

Deutsch, T., Blackwood, D., Eshun, T., & Alexanian, Z. (2017, October 25). "Degree Inflation Hurting Bottom Line of U.S. Firms, Closing Off Economic Opportunity for Millions of Americans."

Deming, David (2017, April), "Increasing College Completion with a Federal Higher Education Matching Grant," The Hamilton Project, Brookings Institution.

Deming, David (2018, February 12), "To increase college completion rates, invest directly in public postsecondary institutions," Brookings Institution.

Doleac, Jennifer (2016, May 31), "'Ban the Box' does more harm than good," Op-Ed, Brookings Institution.

Eggers, W., Hagel, J., & Sanderson, O. (2012, September 21), "Mind the (Skills) Gap," Harvard Business Review.

Executive Office of The President (2015, February), "Economic Report of the President" with the "Annual Report of the Council of Economic Advisors." Transmitted to Congress, Obama Administration.

Flurry, Alan (2017, October 1), "Study estimates U.S. population with felony convictions," UGA Today, University of Georgia.

Gault, Barbara, Reichlin Cruse, Lindsey, Noll, Elizabeth, Sykes, Mary, Talbourdet, Lauren and Eckerson, Eleanor (2016, September), "Child Care for Parents in College: A State-by-State Assessment," Institute for Women's Policy Research.

Guvenen, Faith (2018, February 27), "Stagnation in Lifetime Incomes: An Overview of Trends and Potential Causes," The Hamilton Project, Brookings Institution.

Holzer, Harry (2015, April), "Higher Education and Workforce Policy: Creating More Skilled Workers (and Jobs for Them to Fill)," Brookings Institution.

Holzer, H. (2015, April), "Job Market Polarization and U.S. Worker Skills: A Tale of Two Middles," Brookings Institution.

Institute for Women's Policy Research. Student Parent Success Initiative.

Launch My Career Colorado.

National Center for Education Statistics (2016), "Fast Facts: Distance Learning (Fall 2014)," U.S. Department of Education.

National Employment Law Project, "Ensuring People With Convictions Have a Fair Chance to Work Campaign."

Nedelkoska, L. and G. Quintini (2018, March 8), "Automation, skills use and training," OECD Social, Employment and Migration Working Papers, No. 202, OECD Publishing, Paris.

New Democrat Coalition (2017, August 2), "How Congress Can Help Americans Succeed in the Economy of the Future," Medium.

Opportunity at Work. Who We Are.

Purdue University, Back a Boiler.

Rhode Island DLT, Real Jobs Rhode Island Program.

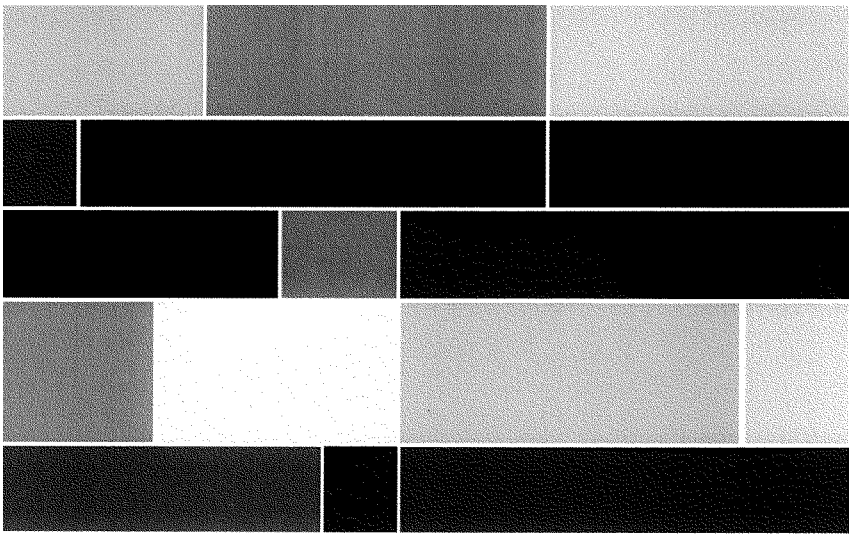
Silver-Greenberg, J., Cowley, S., & Kitroeff, N. (2017, November 18), "When Unpaid Student Loan Bills Mean You Can No Longer Work," The New York Times.

Stephens, R. (2017, April 7), "Automate This: Building the Perfect 21st-Century Worker," Third Way.

Stephens, R. (2017, July 10), "Mind the Gap: The State of Skills in the U.S.," Third Way.

U.S. Government Accountability Office (2011, January), "Multiple Employment and Training Programs.

Whitemore Schanzenbach, Diane; Nunn, Ryan; Bauer, Lauren; Munford, Megan; and Breitwieser, Audrey (2016, October 4), "Seven Facts on Noncognitive Skills from Education to the Labor Market," The Hamilton Project, Brookings Institution.



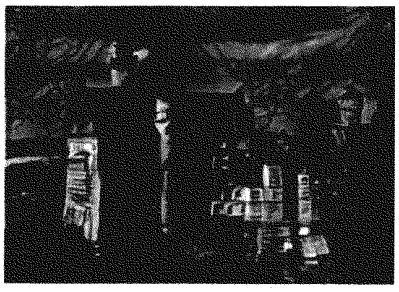
NEW DEMS

REPORT SUBMITTED BY REPRESENTATIVE DUNN

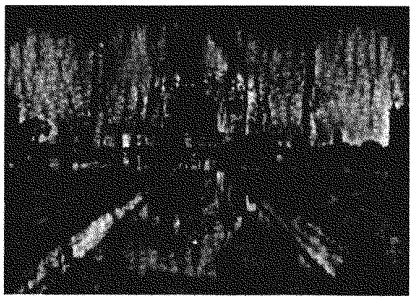
Submission to hearing: "Artificial Intelligence: with Great Power Comes Great Responsibility"
For the Record submitted by Congressman Neal Dunn

Nightmare Machine:

Over the past couple of years, MIT scientists programed a series of three A.I.'s. The first was an A.I. called Nightmare Machine, which was programmed to turn images into haunted imagery: [1]



Neuschwanstein Castle
Schwangau, Germany



Taj Mahal
Agra, India

Shelley:
The second in the trilogy was Shelley; an A.I. Twitter bot. Shelley creates horror stories with the assistance of human collaboration on Twitter:[2]

#Baby

INITIAL STORY BY SHELLEY

When I heard the phone ring again, I ran to the stairs. As I was running down the stairs, I started to hear crying. I shone my phone [1]around the corner of the staircase and saw the crying baby getting closer. I crawled over to it and kicked it as hard as I could. The [2]crying from the stairs turned into a soft metallic sound. [3]

[1] "Nightmare Machine by MIT Media Lab." *Nightmare Machine*, MIT Media Lab, 31 Oct. 2016, nightmare.mit.edu/.

[2] "Shelley by MIT Media Lab." *Shelley*, MIT Media Lab, 31 Oct. 2017, shelley.ai/.

STORY #1 (18 tweets, 6 participants)

When I heard the phone ring again, I ran to the stairs. As I was running down the stairs, I started to hear crying. I shone my phone [1]around the corner of the staircase and saw the crying baby getting closer. I crawled over to it and kicked it as hard as I could. The [2]crying from the stairs turned into a soft metallic sound. [3] I turned back towards the hallway I came from, nothing seemed to be the same, I felt lost, things had been moved from their place [4]and the only thing that caught my attention was the light from the hallway. At this point I could not believe what I was [5]seeing. I turned around and noticed that the furniture had been removed. I had to get some more visibility from the hallway. The crying [6]got louder, I turned around and saw something moving. [7] There it was again,the baby ... he was bloody,but still breathing. And somehow talking to me "Manuel, it is not so easy to do away with me"[8] **"You need to play with me, then I will let you go ... BUT ONLY AFTER I CHEW YOUR FACE"** [9] I stare at the baby and then, for the first time, I realized: The baby... WAS ME! [10]ME! "ME! ME!" I screamed at the top of my lungs. I turned around and saw a face moving back and forth in the baby [11]monitor. [12] Its eyes melted, its face twisted - opening up to reveal a nightmarish scene. I stared, transfixed, but the image flickered. [13]Not even a moment later the camera shut off. [14] I pressed the switch but nothing happened. I look around. In the corner I see something that sends shivers down my spine. [15]It was a man. He was completely naked and white with blood on his hands. He stared at me with those big black eyes, and [16]then he spoke. "I want to show you something." [17] From his wide open mouth, three hands emerged and caught hold of me,dragged me closer to him. Everything went black.[18]

Responses by Shelley are shown in bold.

"Norman":

The third A.I. in this trilogy is "Norman." This A.I. was named after the main character in the famous Alfred Hitchcock movie *Psycho*. Norman's programming was established in the form of data input through image captioning. Image captioning is a deep learning that we as humans often use. For instance, if you show a child a picture of a bird with the caption "bird." The child learns over time that the creature in the image is a bird and keeps that association. This is a learning style also commonly used in language learning software, such as the popular Rosetta Stone program.

The team at MIT used captions from videos on the website Subreddit. Subreddit is a dark website where videos of deaths have been posted for others to watch. In the interest of protecting the scientists working on this project and the families of these victims, the scientists only used the captions from these videos and randomly matched them with Rorschach inkblots. Norman learned these violent scenes through the image captioning learning method [3]:

---

[3] "Norman by MIT Media Lab." *Norman*, MIT Media Lab, 1 Apr. 2018, norman-ai.mit.edu/.

INKBLOT #1
Norman sees:

"A MAN IS ELECTROCUTED AND CATCHES TO DEATH."

INKBLOT #1
Standard AI sees:

"A GROUP OF BIRDS SITTING ON TOP OF A TREE BRANCH."

INKBLOT #2
Norman sees:

"A MAN IS SHOT DEAD."

INKBLOT #2
Standard AI sees:

"A CLOSE UP OF A VASE WITH FLOWERS."

Artificial intelligence is a subject area quickly evolving. The three above cases showcase the evolution of Artificial intelligence in just two short years.

Some of the greatest minds in technology have been quoted to be against the development of artificial intelligence.

Elon Musk: "I think we should be very careful about artificial intelligence. If I had to guess at what our biggest existential threat is, it's probably that. So we need to be very careful,"[4]

Stephen Hawking: "Success in creating effective AI, could be the biggest event in the history of our civilization. Or the worst. We just don't know. So we cannot know if we

[4] Gibbs, Samuel. "Elon Musk: Artificial Intelligence Is Our Biggest Existential Threat." *The Guardian*, Guardian News and Media, 27 Oct. 2014, www.theguardian.com/technology/2014/oct/27/elon-musk-artificial-intelligence-ai-biggest-existential-threat.

will be infinitely helped by AI, or ignored by it and side-lined, or conceivably destroyed by it," [5]

These prompt the questions:

1.) What's next for A.I.? Should we head the warnings of these tech giants?
2.) Will we be able to teach them compassion or empathy? Will we be able to create a perfect consciousness than our own? What will the ramifications of that be?
3.) Is data input stronger than the algorithms that we program the AIs with?

[5] Kharpal, Arjun. "Stephen Hawking Says A.I. Could Be 'Worst Event in the History of Our Civilization'." *CNBC*, CNBC, 6 Nov. 2017, www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html.