

of this writing, Germany²³ and France²⁴ have allocated 3B and 1.5B euros to AI, respectively. The UK²⁵ has pledged an investment of 1B pounds in AI, together with dedicated funding for 1,000 PhDs and 8,000 specialized teachers in AI, and has repurposed its flagship Turing Institute into a major data-driven AI research center. China²⁶ has announced that it will invest billions in AI over the next five years, creating at least four \$50M/year AI Centers and a \$1B/year National AI Research laboratory with thousands of AI researchers and engineers, and committing to training 500 instructors and 5,000 students at major universities. Significant investments in AI, structured with the guidance of this document, will allow the US to take the forefront in propelling the field into a new research era and create significant impact across all sectors of society and the economy.

5. Recommendations

The progress desired in integrated intelligence, meaningful interactions, and self-aware learning pose substantial research challenges that require significant investments that are sustained over time. History tells us that there is no silver bullet, no substitute for sustained effort. For example, speech recognition took four decades of research, mostly in academia, until the work was mature enough to be commercialized (e.g., the commercial Dragon speech recognition software first appeared in the market in the 1990s) and eventually mainstream (e.g., conversational assistants). Similarly, the basic ideas in neural networks were first developed in the 1980s, and it took decades of research as well as improvements in computational resources before they were able to take off. This research Roadmap is designed to prioritize investments and foster similar advances in technologies across AI over the next two decades.

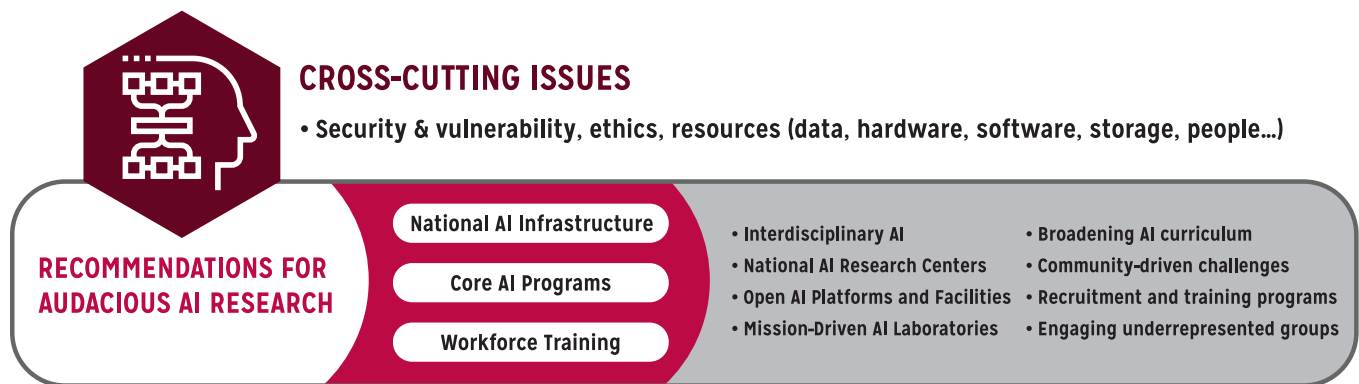


Figure 5. Overview of the 20 Year AI Research Roadmap Recommendations: A national AI infrastructure combined with training of an all-encompassing AI workforce.

²³ "Germany plans 3 billion in AI investment: government paper," Reuters, November 13, 2018, <https://www.reuters.com/article/us-germany-intelligence/germany-plans-3-billion-in-ai-investment-government-paper-idUSKCNII1AP>.

²⁴ "Emmanuel Macron Talks to Wired about France's AI Strategy," Wired, March 31, 2018, <https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/>.

²⁵ "U.K. Unveils \$1.4 Billion Drive Into Artificial Intelligence," Bloomberg, April 25, 2018, <https://www.bloomberg.com/news/articles/2018-04-25/u-k-announces-1-4-billion-drive-into-artificial-intelligence>.

²⁶ "Understanding China's AI Strategy," Center for a New American Security, February 6, 2019, <https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy>.

Achieving the audacious AI research vision put forward in this Roadmap will require a reinvention of the AI research enterprise in order to create a comprehensive national AI infrastructure and to reconceptualize AI workforce training. The rest of this section makes specific recommendations, summarized in Figure 5. A number of important cross-cutting themes underlie all of these recommendations. Progress in AI cannot be decoupled from advances in computational resources and collection of massive datasets, nor can it ignore considerations regarding security, vulnerability, and ethics. As noted in the Code of Conduct of the Association for Computing Machinery, the field of computer science is called to "...Recognize when computer systems are becoming integrated into the infrastructure of society, and adopt an appropriate standard of care for those systems and their user." Ethics must be designed into AI systems from the start. These concerns are particularly acute in the context of high-stakes applications like healthcare, algorithmic bias, and the disparate impact of AI's application on various classes of individuals.

5.1 Create and Operate a National AI Infrastructure

AI as a discipline has reached a level of maturity where significant progress can result from the availability of substantial experimental facilities. The majority of academic AI research in the past has been driven by well-specified tasks in closed domains with short time horizons. Advancing research on AI systems that can take on open-ended tasks, carry out natural interactions, and behave appropriately under changing conditions will require living laboratories that exhibit the dynamics and unpredictability of the real world. Bringing about a new era of audacious AI research will require a National AI Infrastructure that will empower academic researchers while accelerating AI innovation and dissemination of research products.

The recommended National AI Infrastructure is not meant to replace existing AI programs, but rather complement and supplement them with significant resources.

The goal of the National AI Infrastructure is to reinvent the AI research enterprise through four interdependent capabilities: 1) open AI platforms and resources, 2) sustained community-driven challenges, 3) national AI research centers, and 4) national mission-driven AI centers. Figure 6 gives an overview of these components and their interactions. We describe each in turn below.

5.1.1 OPEN AI PLATFORMS AND RESOURCES

There are several key objectives for the Open AI Platforms and Resources:

- ▶ Provide the AI community with substantial experimental resources for both basic research and applications.
- ▶ Attract AI researchers to problems that are harder to tackle without significant startup resources.
- ▶ Promote collaborations across all areas of AI.
- ▶ Support multidisciplinary research by giving non-AI researchers access to AI capabilities.
- ▶ Reduce redundancy of effort and cost in the research enterprise, so research projects do not have to build up capabilities or collect new data from scratch each time.
- ▶ Reduce the cost of individual research programs to integrate relevant capabilities or to compare their work with others.
- ▶ Reduce cost of large teams of collaborators by providing already integrated or easy-to-integrate infrastructure.
- ▶ Provide the AI community with open resources that will bolster not only research in all of academia, but also in small technology companies, companies in other sectors, and government organizations.
- ▶ Provide students with practical training opportunities using operational-grade AI technologies.

- ▶ Attract and retain university faculty with attractive opportunities for research through access to productive ready-to-go environments for unique applications.
- ▶ Promote collaborations between academia, industry, and government in open research environments.

These platforms would include basic components (data repositories, software repositories, knowledge base repositories, software, services, computing and storage resources, and cloud resources), combinations of components into integrated architectures and integrated capabilities, and instrumented facilities (robotic and cyberphysical systems, instrumented homes, simulation environments, game environments, social platforms, instrumented hospitals, etc.). Important attributes of these components include:

- ▶ **AI-Ready:** They should offer capabilities that AI researchers can readily use with reasonable effort. This may require the implementation of services to do basic data preparation, or the integration of datasets with computing resources to facilitate algorithm testing, or APIs that enable new components to be interconnected with existing capabilities.
- ▶ **Open:** They should be shared under standard open licenses (including a range of recognized licenses such as unlimited reuse with attribution, no restrictions, commercial restrictions, share-alike restrictions, etc.) so that anyone can inspect the implementation and characteristics of the component and can reuse and extend it as allowed by the license agreement.
- ▶ **Composable:** They should offer capabilities to facilitate their integration and composition with other components. This involves adoption and development of standards, development of common APIs, mediation through ontologies and public knowledge resources, etc.

There are many efforts to promote community principles and best practices for sharing and reuse of research resources, such as the FAIR principles²⁷, ontology and vocabulary standards (e.g., RDF, Common Logic), provenance standards (e.g., W3C PROV), etc. The AI research community should strive to contribute to the development of best practices and the understanding of what specific requirements are most appropriate to support AI research.

The AI community was one of the first to rely on shared data repositories to drive research. One of the first data repositories ever created was the Irvine Machine Learning Data Repository, which was started in 1987 and was instrumental in the development of effective machine learning methods. The repository still operates today. More recently, many data repositories are associated with competitions and scoreboards (Kaggle, ChiLearn, etc.). For example, ImageNet, a large dataset of labeled images and an associated classification task, helped lead to the 2012 breakthrough in computer vision using deep learning methods. That in turn has helped enable the development of major new technologies, such as self-driving cars. Research in natural language understanding received a similar boost with the creation of the Linguistic Data Consortium (LDC), which was started in 1992 and has made significant investments to create datasets for the natural language community, including datasets with sophisticated annotations such as parse trees for sentences. However, the LDC model relies on a closed-license system, while this Roadmap recommends making AI resources open.

Beyond data, AI researchers have also recognized the importance of other shared resources and community standards. Several robotics platforms have been created over the years for research, experimentation, and commercial purposes, including PR2 (from Willow Garage), AIBO (from Sony), Nao (from Aldebaran Robotics), and Cozmo (by Anki). Software toolkits have also supported more efficient and comparative research, such as ROS in robotics programming, Open CV in computer vision, and Weka in machine learning. Shared communication and representation languages have also promoted integration and comparison of approaches, including KQML, PDDL, and RDF.

²⁷ Forcell. The Fair Data Principles. 2016. <https://www.forcell.org/group/fairgroup/fairprinciples>

Industry labs have made significant contributions to the creation of shared resources for the AI research community. Google's release in 2006 of a dataset of commonly co-occurring words based on automated analysis of more than a billion sentences created a revolution in language technologies. Facebook developed PyTorch as open-source software for deep neural networks, which has allowed the research community to make great strides in machine learning.

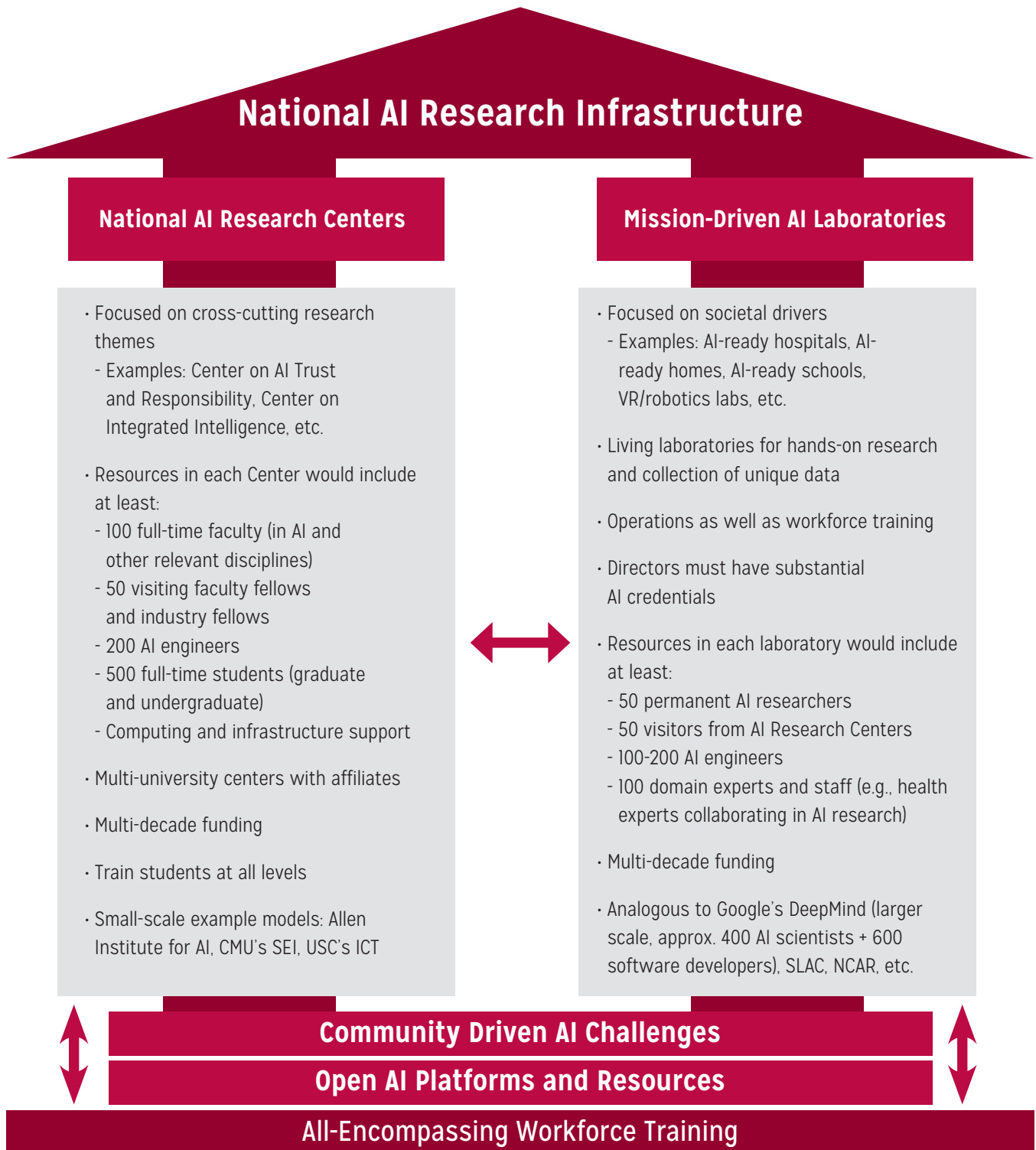


Figure 6. Overview of the envisioned National AI Research Infrastructure.

At the same time, many valuable resources in industry are not publicly shared and are very hard to recreate. Notable examples are the Google Knowledge Graph and AlphaGo. In the case of the Google Knowledge Graph, significant resources were invested to create this proprietary resource. In the case of AlphaGo, the deep net underlying the actual system has never been released, although it would be an amazing resource for AI researchers to study further. It is, of course, understandable that many of these AI resources developed in industry are not shared because they often provide a core competitive advantage for the companies involved.

There is a significant cost associated with maintaining public data repositories in the long term. A general challenge for any research community to sustain publicly available resources is finding a good balance of funding from government agencies, academic institutions, industry support, private donations, and foundation grants.

Unfortunately, open repositories that have been available to the research community in the past have ended up entangled in commercial interests. Figshare started as an open scientific data repository, and once it became successful it was acquired by Springer Nature. Google Open Data was a repository used by many that was shut down in a matter of months. It is important that datasets created by researchers for researchers using public funds be vigilantly protected, and that the Open AI Platforms and Facilities described in this Roadmap be created with necessary provisions for preservation for the public good.

The rest of this section gives an overview of high-priority resources envisioned by the community to be part of the Open AI Platforms and Resources, which include AI-ready data repositories, open AI software, open AI integration frameworks, an open knowledge network, and testbeds. The next sections describe each of these kinds of resources in detail.

AI-Ready Data Repositories

Because the collection and preparation of data often requires a significant amount of effort, the public availability of well-designed datasets are important drivers of research progress. There are many useful principles in the design of these data repositories that currently are not, but should be, widely adopted—such as proper documentation, quality control, clear problem and task statements, and sometimes incentives such as awards and scoreboards.

However, available data repositories used in the AI community tend to focus on very few tasks and very few data formats, because they are:

- Largely made up of labeled data for classification problems. Unfortunately, classification is a very narrow task within the much broader scope of capabilities that AI systems need.
- Fit for traditional formats that are easy to characterize, use, and share. These tend to be curated compilations of images, text, and other readily available data. This is a very particular kind of data that lends itself to simple tasks (e.g., classification).
- Designed to demonstrate utility of algorithms, rather than utility of solutions.
- Collected for applications of commercial or specific research interest. Very few shared datasets are of public or societal interest.

For example, there are many data repositories for classification of text documents, or images of pets. There are few with sensor data for autonomous cars, personal task management (e.g., email exchanges, to do lists.), social interactions, etc. This makes it more challenging to address the forward-looking AI research areas outlined in this Roadmap.

Another important issue is that the many other shared data repositories are not easily usable by AI researchers, notably scientific and government datasets. Many scientific datasets would be appropriate for AI research, but are not easily findable outside the particular scientific community or are not easily usable by domain experts in the particular field of science. In some cases, scientific datasets are specifically prepared and intended for use outside the science domain but are not well known to AI researchers. Open government data offers another excellent potential source of data and concomitant opportunities, but the data are not in formats that are well documented or accessible.

Well-designed AI-ready dataset repositories will have a number of benefits:

- ▶ Allow researchers to seek and share datasets that address real-world problems and tasks, particularly in sectors and applications concerned with public and societal interests.
- ▶ Increase the opportunities for and productivity of AI researchers by making it easier to use datasets in their research, and facilitate comparing their solutions to others.
- ▶ Promote datasets that are available but not yet in a form that is conducive to research, such as government data and scientific datasets.
- ▶ Support data sharing for domains and tasks that have complex data, such as social communication, interacting behaviors, multi-criteria optimization, real-world goal accomplishment, etc.
- ▶ Promote research on managing shared datasets that contain sensitive data (e.g., personal data).

AI Software and AI Integration Frameworks

Open software facilitates research by providing a significant substrate to build upon. Some software libraries have come from academia (OpenCV, NLTK), others from industry (TensorFlow), and others through a combination of both (ROS). The availability of software that implements powerful approaches in an easily accessible manner spurs significant innovation. The open software movement provides many examples of the benefits of organic collaborative communities to support, and in many cases extend, software to production-grade quality.

AI needs an open collaborative ecosystem that includes a range of governance strategies for the software. Open software can be shared with minimal oversight through its publication with open licenses in open repositories with version control, a common practice today that will no doubt continue to be widespread. An alternative approach to open software is to form well-organized communities and processes organized around key software projects. For example, the framework provided by the Apache Software Foundation supports incubation of communities and establishment of governance and decision making, and has been a very successful approach to developing robust operational-grade software for both research and deployments.

Beyond software libraries, AI integration frameworks are a high priority in this research Roadmap. A wide range of approaches is possible, with varying benefits and tradeoffs. On one side of the spectrum, frameworks for integrating AI components would rest on the existence of interchange standards and the practical use of modularity and composability principles. Composing AI capabilities is challenging because they may have interacting behaviors and different assumptions that would need to be harmonized. The more heterogeneous the capabilities, the more challenging their integration. On the other side of the spectrum are frameworks that are designed in principled ways and can be adopted to develop components and capabilities that are well integrated by design. Both approaches can lead to a range of designs, depending on task constraints. For example, approaches may be based on cognitive science findings to create more human-like systems, or engineered for specific types of applications, or designed to support real-time robot platforms.

The availability of open AI integration frameworks and environments would encourage research on AI systems that combine a wide range of tasks and capabilities. This would also encourage the exploration of a diverse range of approaches and their assessment for different uses and purposes.

5.1.1.1 An Open Knowledge Network

A major theme in the research agenda outlined in this Roadmap is the need for AI systems to access significant amounts of knowledge about the world. The AI community has a tradition of knowledge sharing, promoted through standards (such as KQML, RDF, OWL), shared ontologies and knowledge bases (e.g., OpenCyc, WikiData), and open libraries and tools (Protégé, etc.). These traditions and practices must be complemented by a push toward generating large-scale knowledge bases, capturing both a wide range of everyday knowledge and specialist knowledge in important application areas (e.g., health, science, engineering).

Knowledge repositories have become a key competitive advantage for AI applications in industry. As tech giants push the envelope in markets such as search, question-answering, entertainment, and advertising, they have discovered the power of knowledge to make their systems ever more capable. These knowledge sources are usually captured as a graph containing entities of interest (e.g., products, organizations, notable people, events, locations, etc.) interlinked in diverse dimensions, and are highly structured to be machine-readable. Immense resources are being invested in Google's Knowledge Graph, Microsoft's Satori knowledge graph, Amazon's Product Graph, Facebook's Social Graph, and IBM Watson's Knowledge Graph. In addition to the big technology companies, many other large companies are creating knowledge graphs to drive innovative products (e.g., Thompson Reuters, Springer Nature, LinkedIn).

Although many of the technologies underlying these knowledge bases were originally developed in academia, researchers outside proprietary boundaries have limited or no access to these resources and no means to develop equivalent ones. For example, recent accounts of the Google Knowledge Graph place its size at 1 billion entities and 70 billion assertions about them, and Microsoft's at 2 billion entities and 55 billion assertions. These represent major investments that are very hard to replicate, but are all proprietary and not accessible to outside researchers. In addition, these knowledge bases are developed to support the commercial interests of the companies that created them, such as search results and ad placement. They do not necessarily contain material relevant to scientifically vital domains and other applications. Nor do they contain the broad commonsense knowledge that provides a useful substrate for the application of professional knowledge to daily life, common ground for communicating with people, and knowledge of the physical and social world needed by robots and other AI systems.

An *Open Knowledge Network (OKN)* with very large amounts of knowledge about entities, concepts, and relationships in the world would foster the next wave of knowledge-powered artificial intelligence innovations with transformative impact ranging from scientific research to the commercial sector. The OKN would enable new research in personal assistants who could better understand the world in which they operate, machine learning systems to enrich data with extensive information about the underlying objects, natural language systems to link words and sentences to meaningful descriptions, and robotics devices with context and world knowledge—among many others.

A recent report describes several key characteristics that the community considers desirable of a public resource such as the OKN.²⁸ First, it should be a highly distributed network of knowledge resources, with open contributions and open access to allow a diversity of uses. Second, the underlying representation should be extensible as further research is done on languages that accommodate more complex forms of knowledge. Third, the provenance of each piece of knowledge (i.e., its origins), should be explicitly represented. Fourth, the context where each piece of knowledge may be true should be qualified, whether temporal (e.g., a person is a college student only during a few years), spatial (e.g., seagulls are a common bird in coastal areas but not inland), etc.

The availability of the OKN would enable basic research in advancing the underlying technologies in order to provide AI systems with extensive knowledge about the world. Further research is needed to extend the paradigm of entities and their relationships, which is only sufficient to capture very basic knowledge. Much research is also needed in facilitating the variety of uses and application approaches that will be necessary, for example to answer new kinds of questions and to provide natural interfaces to users. New approaches to extract knowledge from text, images, videos, and interactions with people will be needed to keep the OKN up to date and to rapidly extend it for new applications and domains of interest.

The OKN should also include efforts to build knowledge resources for domains of societal interest such as science, engineering, finance, and education. Scientific research could use the OKN as a gateway to a vast trove of knowledge and data, and particularly to disseminate knowledge and data across disciplines.

²⁸ Networking and Information Technology Research and Development Program. Open Knowledge Network: summary of the Big Data IWG Workshop. Oct. 4-5, 2017. <https://www.nitrd.gov/pubs/Open-Knowledge-Network-Workshop-Report-2018.pdf>

The OKN would also give industry and government organizations access to extensive world knowledge. Today, only the biggest tech companies have the resources to develop significant knowledge graphs. With an open resource such as the OKN, all companies, regardless of size and sector, would be able to use world knowledge to develop AI systems that can operate effectively in that world.

5.1.1.2 AI Testbeds

AI testbeds must strive to balance tractability and real-world relevance. Many researchers choose to study simplified tasks in closed domains, rather than open-ended real-world problems, because toy tasks are more tractable for today's methods. In order to draw researcher attention, a task should admit partial progress with today's algorithms and architectures, yet not be fully solved. At the same time, overly simplified testbeds can drive illusory progress that appears impressive in the lab but does not transfer to the real world. In order to strike this balance between research tractability and real-world relevance, testbeds need to be carefully and iteratively crafted in a tight feedback loop between testbed designers and AI researchers.

AI testbeds would also help establish and maintain benchmarking standards to promote measurable research progress, allowing for principled comparisons of different approaches and facilitating reproducibility, thereby enabling the research community to make significant measurable progress.

The creation of a successful testbed should be rewarded as a valuable scientific contribution. New approaches that attain better performance on pre-defined well-scoped problems are often rewarded in the current system of publication incentives. By contrast, applications of existing algorithms to new domains, including the construction of novel testbeds that capture important real-world dynamics, are often viewed as not innovative enough for a publication in top conferences and journals. New incentives must be created to reward well-designed testbeds as valid contributions to research.

Best practices and guiding principles of AI testbed design will need to be developed. There are important research directions in determining whether a testbed captures a particular characteristic of a real-world environment, summarizing how the observations collected differ from real data, and whether those differences will affect the validity of AI systems developed on the testbed to perform well on the real environment.

The rest of this section gives examples of the kinds of testbeds that would propel AI research areas outlined in this Roadmap.

5.1.1.2.1 Human-Machine Interaction Testbeds

Today, most of the progress on AI technologies is on simple perceptual tasks that require reasoning over a single interaction modality, e.g., images or text. However, many crucial real-world tasks require systems that can handle and reason over information available from multiple modalities and information sources. Furthermore, most successful AI techniques are data-driven; to build such AI systems we therefore need to develop and make widely available high-quality research ecosystems such as large datasets and software libraries. Much of the progress in AI subfields such as computer vision and natural language processing has been made possible due to openly available datasets (such as ImageNet, Gigaword Corpus, SQUAD, etc.) and software libraries (such as Tensorflow, PyTorch, AllenNLP, etc.). Similarly in robotics, the open-source Robot Operating System (ROS) has become an international standard with a massive international user and contributor community that includes academia, industry, and volunteers. Tackling future interaction research challenges requires creating large datasets that represent real-world problems and developing open-source software libraries to enable rapid prototyping and deployment at scale. Developing software libraries is especially important due to the sheer complexity and the engineering effort required to handle streams of multimodal data from many sensors.

The availability of rich interaction research platforms would enable new research to address significant real-world problems concerning human communication and human-machine interaction. They would enable the collection of data to cover the diversity and range of human communication. This diversity is not only concerned with conversation topics or communication modalities and channels, but with people of different ages (children vs. elders), different native languages, cognitive and physical abilities, speech impediments, background and technical skills, personal goals and priorities, and other attributes that affect their communication needs for AI systems. Including such diversity is crucial for generating AI systems that can operate robustly in the real world.

Another barrier to human-machine interaction research is the lack of data from realistic, and often sensitive interactions, such as those with vulnerable populations (children, elderly), private settings (doctor's office, therapist session), and group contexts (at work, school, on public transportation, etc.). The development of intelligent systems that can operate in such settings and with a diversity of users requires training data, yet such data currently take years to obtain, and are only available to a small group of researchers, due to the cost, lengthy process, and privacy constraints, which collectively inadvertently result in slow progress. Large datasets from such settings, obtained with proper consent and shared under appropriate privacy rules, would serve as major enablers for AI research into human-machine interaction and other areas of AI.

5.1.1.2.2 Life-Long and Life-Wide Personal Assistant Platforms

An exciting application area for AI with potential for high impact in society is life-long and life-wide personal assistants (LPAs). Today's personal assistant technologies are very limited, and having an open ecosystem for research in this area could lead to significant advances and have significant impact in society. Needless to say, ethical considerations (privacy, security, and accountability) are a primary concern in the design of these systems.

LPAs are AI systems that assist people in their daily lives at work and home, for leisure and for learning, for personal tasks and for social interactions. LPAs are life-long in that they will assist a person for long periods of time, possibly for their lifetime, remembering events and people that a person may not naturally recall well and adapting to a person's needs as their life changes incrementally (new preference, new daily habits) or drastically (new job, new interest). LPAs are life-wide, operating across all areas of their owners' lives to build and maintain context-sensitive models of those users.

LPAs will not only be personalized, but also personal, in that all data shared with them will remain the property of the agent's owner. Current personalization methods typically offer minor language adjustments and content recommendations based on expressed and learned preferences. In the future, personal AI systems will learn how to best serve the interests of the people interacting with them. They will learn what sort of language is most motivating, which feedback improves a person's well-being, and which contexts are most engaging for the owner. LPAs will reason about social context and the social consequences of actions. They will also support human-machine teaming and automate tasks to achieve effective collaboration.

LPAs will learn and adapt over their users' lifetime. They will detect situations for opportunistic training and education based on modeling the user, detecting opportunities, and creating curriculum on the fly. This learning will support users both professionally and personally, serving as a cognitive multiplier for increased efficiency and learning, enhancing personal autonomy, assisting people in their daily lives, and empowering users to overcome mental and physical limitations.

LPA technology has the potential to be a game-changer across multiple sectors, including smart cities, psychology, social behavior, human body and brain function, education, workforce training and re-training, eldercare, and health. LPAs could oversee therapy, coaching, and motivation that supplement human care, and can continue long-term therapy in the home, both after hospitalization and for chronic conditions. LPAs could help people with cognitive tasks, including supporting long-term and short-term knowledge and memory.

An open research ecosystem for the development of LPAs would propel research as well as technological and commercial innovation in this area. This vision of LPAs requires innovations that cannot happen unless researchers have the ability to access and analyze data and to carry out experiments to test new approaches. Currently, personal assistants are developed and operated in the private sector, so the collected data is proprietary and the underlying platforms are not accessible for academic research. Furthermore, important privacy issues need to be addressed that can better take place in an academic open setting. Almost every AI research challenge discussed in this document is relevant to the development of LPAs. Their overall design is also a fundamental research challenge, since an LPA could be a single AI system with a diverse range of capabilities or a collection independently developed specialized systems with coordination mechanisms for sharing user history and personalization knowledge. Open platforms for continuous data collection and dynamic experimentation would significantly accelerate research, not only in this important application of AI but also in basic AI research.

5.1.1.2.3 Robotics and Physical Simulation Testbeds

Physical robot development platforms are unaffordable and inaccessible to many, and simulations are only possible and available for some types of platforms and AI functions. In order to facilitate research on integrating intelligent capabilities into robots and developing robust real-world systems, accessible robotics testbeds must be created that allow experimentation and testing that bring together diverse AI components, including perception and fusion from multiple sensors, speech processing and natural language interfaces, reasoning and planning, manipulating objects in the world, navigation, interacting safely and naturally with people, and accomplishing a variety of tasks and goals. Open robotics testbeds will be critical enablers for making progress in combining AI capabilities and robotics in real-world problem contexts. To enable progress toward real-world robot deployment, robotics testbeds will also need to be embedded in realistic environments: instrumented homes, schools, and hospitals, advanced manufacturing facilities, public venues such as museums and city centers, and large-scale environments such as farms. These environments should be equipped with ambient sensors (high-fidelity cameras, microphones, etc.) and Internet of Things infrastructure, as well as virtual reality and augmented reality interfaces, to support diverse types of interactions with people. They should contain various robot platforms appropriate for their particular contexts (mobile manipulators, table-top robots, swarm-type robots, high-degrees-of-freedom arms, humanoid robots, field robots). The physical environments are critical for real-world robot development; however, testbed software infrastructure will need to include high-quality simulations to facilitate ramping up and scaling projects.

These robotics testbeds will require privacy-protected data storage and data processing facilities. They will also need to provide high-speed communications to allow for performing real-time experiments remotely. They would need to be supported with multiple technical staff who facilitate research and data collection and maintain the hardware and software infrastructure.

In addition to physical environments for robotics, AI research will also greatly benefit from high-quality, realistic virtual environments for training robotic systems. For many kinds of robotics problems, robot learning methods require significant training time (on the order of months or years of experiment time), which can only be generated in virtual environments. The creation of virtual training environments that provide a sufficiently detailed and realistic replica of the physical world is a major engineering undertaking and a critical enabler for physical machine intelligence.

5.1.2 SUSTAINED COMMUNITY-DRIVEN AI CHALLENGES

AI challenges with million-dollar rewards, such as the X Prize, the DARPA Robotics Challenge, and industry-driven competitions, attract significant attention and catalyze creative solutions. However, they lead to idiosyncratic non-generalizable approaches, rather than the kinds of profound new ideas that lead to transformative research in AI. Great excitement is also generated by AI competitions with common datasets and scoreboards, but they tend to target surgically delineated and unconnected problems. To move AI to the next step, we need to capitalize on the energies fostered by healthy competition, while promoting concerted progress in hard AI problems.

To that end, we recommend creating a set of sustained community-driven AI challenges with the following important features:

- **Research-driven:** The competitions would be created to drive AI research based on an explicitly articulated research Roadmap with milestones along the way.
- **Community governance:** They would involve organizational structures that enable AI researchers (rather than outsiders) to agree to the definition of a challenge and evolve the rules of the competition as research progresses.
- **Grassroots participation:** Challenges and competition rules would be designed by competition participants collaboratively.
- **Regular updates:** Competition rules would be updated based on research progress in ways that direct the community toward new advances.

- ▶ **Decomposable:** Ambitious challenges could be broken down into more tractable problems over a sensible timeline to formulate a feasible long-term research arch.
- ▶ **Resource creation:** Participants should be supported and incentivized to create shared resources, useful tools, and infrastructure that supports community participation in the competitions.
- ▶ **Embedded metrics:** The community should develop the necessary metrics to measure research progress based on the expertise gained by analyzing competition entries.
- ▶ **Compete-collaborate cycle:** Competing teams should gather after a competition to share their approaches and open-source software, to collaborate in updating the competition rules, and to regroup to undertake the next competition.
- ▶ **AI ethics:** These important considerations should be incorporated into the competitions at all levels.

A good model for sustained community-driven AI challenges is the RoboCup Federation. Started in 1993 to use the game of soccer to drive AI and robotics research, its ultimate goal is: “By the middle of the 21st century, a team of fully autonomous humanoid robot soccer players shall win a soccer game, complying with the official rules of FIFA, against the winner of the most recent World Cup.”²⁹ However, it is acknowledged by the Federation that this is a challenging long-term goal, and that other research and discovery will occur along the way that will have significant societal impact. At the moment, in addition to RoboCupSoccer the federation runs other leagues under the rubrics of RoboCupRescue, RoboCup@Home, and RoboCup Industrial. The RoboCup community includes more than 1,000 people. The competition rules for each league are defined by a committee elected by the participants in that league. The rules for the competitions are designed to push the research along dimensions proposed by researchers participating in the league based on their best intuition about promising directions and anticipated topics of their funded research projects. The rules are updated each year as needed to push the research. Annual research symposia enable competition participants to share the major ideas and advances that enabled them to showcase new capabilities at the competitions. Many teams collaborate to develop research infrastructure, such as simulators or program interfaces. Companies have developed affordable platforms to support the competitions, and continue to be involved in proposing new RoboCup leagues and research challenges of interest.

The recent growth of AI for social good presents an opportunity for pursuing a sizeable variety of AI problems with the unifying theme of contributing to solving societal challenges, including human health, safety, privacy, wellness, sustainability, energy use, and many others. Applying the RoboCup model to the social good context would create a motivating opportunity for numerous AI researchers as well as students to contribute to meaningful problems.

5.1.3 NATIONAL AI RESEARCH CENTERS

The National AI Research Centers are intended to create unique and stable environments for large multidisciplinary and multi-university teams devoted to long-term AI research and its integration with education. While the majority of current academic research tends to be done with precisely focused, short-term, and single investigator projects that support two to three faculty, these centers will provide the expertise and critical mass needed to make significant advances in otherwise unattainable AI research goals.

Each National AI Research Center would be a catalyzer for a broad and substantially challenging AI theme that would serve as a pivot for the center’s research. This research Roadmap suggests many priority areas that could serve as pivots, such as open knowledge networks, cognitive architectures, fair and trustworthy AI, representation learning, assistive robotics, LPAs, and many more.

²⁹ RoboCup. Objectives. <https://www.robocup.org/objective>

National AI Research Centers are envisioned to be:

- ▶ Funded through decade-long commitments, to provide stability and continuity of research
- ▶ Multi-university centers with a smaller core set of partners and a large network of affiliated educational institutions, including research universities, liberal arts colleges, community colleges, and K-12 schools.
- ▶ Multidisciplinary in the expertise involved in the research areas
- ▶ Multi-faceted in their research goals
- ▶ Inclusive hosts for visitors bringing diverse views and new ideas
- ▶ Effective dissemination vehicles for significant results

These Centers would provide:

- ▶ Practical environments for the development of open resources
- ▶ Governance and oversight for relevant Community-Driven AI Challenges
- ▶ A fertile ground for involving other in other areas of computer science and other disciplines in AI research
- ▶ Unique educational opportunities for training the next generation of AI researchers, engineers, and practitioners

Each National AI Research Center would be funded in the range of \$100M/year for at least 10 years. With this level of funding, such a center would be able to support an ecosystem of roughly 100 full-time faculty (in AI and other relevant disciplines from different universities), 50 visiting fellows (faculty and industry), 200 AI engineers, and 500 students (graduate and undergraduate), and sufficient computing and infrastructure support.

There are a few examples of AI research centers that have long-term funding. The University of Maryland's Center for the Study of Language (CASL), founded in 2003 as a DoD-sponsored University Affiliated Research Center (UARC) funded by the National Security Agency, includes about 60 researchers and 70 visitors from academia and industry focused on natural language research with a defense focus. The Institute for Creative Technologies at the University of Southern California, established in 1999 as a DoD-sponsored UARC working in collaboration with the US Army Research Laboratory, has about 20 senior researchers, 30 students, and 50 support staff working on AI system and virtual environments for simulation and training. The Allen Institute for AI, which was established in 2014 and is privately funded, has more than 40 senior researchers and 60 support staff focused on machine reading and commonsense reasoning. *Although these centers have long-term funding and support staff, all are at significantly smaller scale than the ones envisioned here, and none are multi-institutional centers.*

5.1.4 MISSION-DRIVEN AI LABORATORIES

The Mission-Driven AI Laboratories (MAILs) are intended to be living laboratories for AI development targeting areas of great potential for societal impact with high payoff, such as healthcare, education, and science. While many aspects of academic research are driven by applications of societal interest, current work tends to be piecemeal and does not always target deployment.

MAILs are envisioned as living laboratories that would serve as real-world settings for:

- ▶ Unique collaborations between AI researchers, domain experts, and practitioners to formulate challenges and requirements for AI technologies
- ▶ Immersive environments for multidisciplinary teams to study challenging practical problems
- ▶ Exposing AI researchers to hands-on experience, allowing them to attain deep knowledge about challenging societal problems

- Collection of unique research-grade comprehensive data pertaining to challenging real-world problems
- Testing of AI prototypes for formative evaluation
- Experimentation with novel processes and approaches for the deployment of new AI capabilities
- Enabling industry and academic researchers to work in open environments on problems of mutual interest
- Testing the safety, security, and reliability of AI systems before broad adoption
- Training the next generation of AI researchers, engineers, and practitioners
- Exposing students and researchers from other fields to advanced AI technologies relevant to their domain and in a practical setting

MAILs would:

- Work closely with the National AI Research Centers to integrate and assess their most recent advances
- Provide sustained infrastructure, including AI engineers and facilities with associated personnel to support AI research
- Stress test the open AI platforms and resources
- Provide environments and data to formulate community-driven AI challenges
- Demonstrate the operational value and ethical integrity of AI technologies and advances
- Work closely with the National AI Research Centers to integrate and assess their most recent advances

Each MAIL would fund AI research in the range of \$100M/year for several decades, with a separate budget for operations and support of the center. This would be a reasonable period of time to see significant returns on investment and transformative research in the target areas. With this level of funding, a MAIL could support an ecosystem of roughly 50 permanent AI researchers, 50 visitors from the National AI Research Centers at any given time, 100-200 AI engineers, and 100 domain experts and staff focused on supporting AI research. MAILs should be run by people with substantial AI experience and credentials in order to bring together the research community and ensure that research quality remains a priority.

There are no existing examples of mission-centered AI laboratories. Outside of AI, there are many such experimental laboratories in other sciences that could serve as models. The SLAC National Accelerator Laboratory, founded in 1962 as the Stanford Linear Accelerator Center, has led to four Nobel-winning results in particle physics to date. Though initially staffed by experimental physicists, it eventually became an academic department at Stanford and a national Department of Energy laboratory. SLAC is home to more than 1,500 researchers, support staff, and visitors from several hundred universities in 55 countries from both academia and industry. Its annual budget is over \$430M. Other DoE laboratories are also associated with universities, including the Oak Ridge Laboratory offering an interesting model with the Oak Ridge Associated Universities (ORAU) with more than 100 affiliated universities across the US. A different model is the National Center for Atmospheric Research (NCAR) is a Federally Funded Research and Development Center (FFRDC) through National Science Foundation grants focused on atmospheric and space sciences and their impact on the environment and society. NCAR is run by the University Corporation for Atmospheric Research (UCAR), a nonprofit consortium formed by more than 100 universities. NCAR is home to more than 700 researchers and support personnel, and its annual funding surpasses \$165M. Another model could be the Jet Propulsion Laboratory (JPL), a FFRDC funded through NASA awards that is administered by the California Institute of Technology with a mission focus on space and Earth sciences. Its annual budget is near \$1.5B, supporting more than 6,000 employees. Interestingly, JPL has traditionally had a substantial number first-rate AI researchers developing and deploying award-winning AI technologies in space missions, with a concentration on robotics, autonomous planning and scheduling, and machine learning.

The following sections illustrate some potential target mission areas for MAILs, including AI-ready hospitals, AI-ready homes, AI-ready schools, and experimental facilities for preparation and response for natural disasters.

AI-Ready Hospitals

AI systems have the potential for tremendous impact in improving care in hospitals, including administrative support (e.g., bed management and staffing), care delivery (e.g., monitoring for treatment administration and early warnings of potential deterioration, assisting with integrated medical information) and decision making. Today, researchers face significant challenges in investigating effective use of AI methods in these arenas, because doing so requires access to current patient populations, not simply working with retrospective datasets. It requires significant time and effort to develop the collaborative research efforts with clinicians and build the cross-disciplinary understandings needed to make progress in these areas. The constraints on access to health data and on testing, as well as the taking of clinician practice time, make exploratory collaborations in a real setting difficult. Although this is an area of high potential impact and great interest to AI researchers, to date it has been difficult to move beyond retrospective datasets that lead to pilot studies, with full trials rare and typically only after many years.

AI-ready hospitals offer one possibility for making such research more effective. These hospitals would be designed to enable AI researchers to become part of a hospital ecosystem and easily collaborate with clinicians and hospital managers to investigate the development of novel uses of AI in healthcare settings. The IT systems for these hospitals would be designed in collaboration with AI researchers to enable the collection of rich data and easy prototyping. They would collect appropriate information and support AI systems' use of that information (both adding to records and obtaining information from them). Hospital personnel, including IT engineers and healthcare professionals, would be available to collaborate with AI researchers to take care-flow constraints into account, to understand data collection and data quality, to ensure proper access of data to respect privacy, and to design appropriate user interfaces.

We recognize that the development of such hospitals raise a wide range of ethical and societal challenges related to patient participation (could a patient opt out?), location choice (could such a hospital be the only hospital in some urban neighborhood or rural area?) and clinician participation, as well as the full spectrum of privacy issues. AI-ready hospitals or healthcare systems that enable (relatively) rapid development and testing of AI approaches could be created in diverse locations, to help ensure robustness across patient populations and environments. These could be set up as learning hospitals, where medical and nursing students would also acquire skills in advanced technologies for healthcare. In such a context, there would be a staff of clinicians, clinical trial specialists, and IT engineers whose mission is to team with AI researchers for problem selection, experimental design, data pulls, and selective access only to patients who agree to be part of a given study. This effort would be supported by HIPAA-compliant computational resources.

AI-Ready Homes

AI-ready homes could provide a testbed for research on AI systems that support people aging in place, improving the safety and security of homes, and making life more convenient. These heavily instrumented homes would continually provide data to researchers in agreed-upon spaces and times in and around the house. Privacy protocols would be in place governing any researcher-access to the data, including the use of identity scrubbing (e.g., automatic replacement of faces with abstract depictions, text transcripts of speech instead of audio) where feasible. Security measures would also be designed into these systems to assure their occupants appropriate control over their operations.

AI-Ready Schools

Much of the best work in AI for education has been done in close collaboration with in-service teachers, but establishing such collaborations is difficult, given the other demands on teachers' time and the organizational constraints of schools. AI-ready schools would encourage participation in such collaborations as one component of a teacher's job, supported and rewarded by the schools. AI-ready schools would have agreements in place for their students to participate in experiments with AI software.

Such schools would be provisioned with IT infrastructure to support a range of experimental AI systems, including machines that could be loaned to students for homework and servers for local data collection. Data gathering and processing would comply with the Family Educational Rights and Privacy Act (FERPA) using, for example, on-site anonymization, and would support educational data mining and experiment analysis without compromising student or teacher privacy. These schools would be distributed in regions that serve diverse populations, in order to ensure that the AI systems developed are robust across the full spectrum of students who need them.

AI-Ready Science Laboratories

Already many laboratories use software to control experimental apparatus and store data in digital form. AI-ready laboratories would go beyond this by including infrastructure for interfacing instrumentation in that field with AI systems, thereby producing data with built-in semantic description. This would enable automated analysis and experimentation and facilitate sharing and synthesis of data across experiments and laboratories. Capturing thinking and decision making processes with multimodal sensors, and enabling AI systems to participate in these conversations, would also be supported. Protocols governing data sharing that respect the privacy and priority of the domain scientists, as well as the AI scientists, would be in place.

AI for Natural Disasters and Extreme Events

AI can be used in many aspects of preparation for and response to natural disasters and extreme events, such as hurricane winds and storm-related flooding, which is predicted to affect 200 million Americans across 25 states, causing as much as \$54B in annual economic losses.³⁰ Before such an event occurs, AI could be used for improving predictions, integrating geosciences models, analyzing population and asset risk, and designing robust plans in preparation for extreme events. After the onset of such an event, AI can be used in a variety of ways, from coordination and prioritization of tasks for first responders, dissemination of information and control of misinformation among the affected population, and search and rescue (e.g., via rescue robotics).

5.2 Re-Conceptualize and Train an AI-Encompassing Workforce

Comprehensive changes need to be undertaken in order to restructure and train a diverse AI workforce to prepare highly skilled researchers and innovators. Although the demand for AI expertise is very high, the particular expertise required for any job will vary widely in terms of topics, technical skill, and experience. For example, in some cases advanced knowledge of AI algorithms will be required; in others, the demand will be for practical data analysis and user experience design skills, in still others, software engineering will be of most value. Because AI systems and supporting technologies will evolve very rapidly, training AI engineers to continuously adapt to new techniques and tools will be essential. Equally essential will be a strong focus on ethical issues around AI, including safety, privacy, control, reliability, and trust, throughout every level of this effort, in order to instill ethical thinking in all AI professionals.

This section discusses eight major recommendations: 1) Developing AI curricula at all levels, 2) Recruitment and retention programs for advanced AI degrees, 3) Engaging underrepresented and underprivileged groups, 4) Incentivizing emerging interdisciplinary AI areas, 5) AI ethics and policy 6) AI and the future of work, 7) Training highly skilled AI engineers and technicians 8) Workforce retraining. Success at all of these levels will be key for providing the general public and decision makers with general knowledge about AI that will be fundamental for everyone, both as intentional consumers or unintentional recipients of AI technologies. Professionals, as well as the public, should be informed about the high stakes of the development of AI and educated about appropriate expectations and implications.

³⁰ 30 Congressional Budget Office, Expected Costs of Damage from Hurricane Winds and Storm-Related Flooding, April 10, 2019. <https://www.cbo.gov/system/files/2019-04/55019-ExpectedCostsFromWindStorm.pdf>

5.2.1 DEVELOPMENT OF AI CURRICULA AT ALL LEVELS

We recommend comprehensive curricula for K-12, undergraduate, graduate, and postgraduate AI education. AI curricula guidelines should be developed to start at an early age to encourage interest in AI, understanding of the associated issues and implications, and curiosity to pursue careers in the field. AI curricula materials should be made freely available and teacher-training opportunities should be supported and promoted. Curricula at all levels should contain a significant part dedicated to ethical issues around AI, including safety, privacy, control, reliability, trust, etc. These issues should be integrated into the computer science curricula as a whole, rather than in separate standalone courses. To facilitate operationalization, opportunities, support, and incentives should be provided for current K-12 teachers to be trained in the use of AI curricula and course materials.

We recommend guidelines for undergraduate, graduate, and professional courses and degrees and certificates in AI, which educational institutions at all levels can employ in order to set up programs that prepare students for careers in the field. This should go well beyond computer science and engineering majors, so that programs are available to all undergraduate students to incorporate AI minors or AI concentrations as complements to their fields of study. AI undergraduate courses should not all be highly technical, so that offerings and learning opportunities are made available to and accessible for all students in all majors across a campus. Here, too, coverage of ethical issues will need to play an important role.

We recommend graduate courses and degrees that provide students with opportunities to learn about all relevant aspects of AI, particularly in practical settings, in order to preserve US leadership in AI. Courses developed jointly by faculty in different areas often lead to creative research outcomes and in many cases change the interests and career direction of students. Creating a heterogeneous ecosystem of AI and AI-related graduate studies will effectively promote much-needed spread of ideas across fields that will greatly benefit AI research and innovation. The associated course materials should be openly shared to complement curriculum guidelines, allowing every institution to offer a wide variety of AI courses. More recommendations regarding interdisciplinary studies are included below.

We also recommend developing guidelines for professional programs and certifications that will appropriately qualify individual expertise and experience in all aspects of AI, including ethics. This will enable a wide range of organizations to provide the AI education opportunities that will be needed to meet the demand for professional careers and retraining in AI. These certifications will be particularly important for jobs that require high-stakes decision making and for organizations where AI expertise is not already available.

5.2.2 RECRUITMENT AND RETENTION PROGRAMS FOR ADVANCED AI DEGREES

We recommend significant increases in the production of graduates with advanced degrees in AI. First, the AI education and training programs proposed in this report cannot be staffed unless significant numbers of faculty, teachers, and instructors are available with advanced AI degrees. Second, significant AI breakthroughs and innovations such as those outlined in the research plan of this report will be accomplished at speeds and quantities that are proportional to the availability of AI researchers with advanced degrees. Third, the more advanced AI expertise that is available to permeate different disciplines and areas of society, the more transformative AI technologies will be—across the board.

The creation of the AI centers described in the previous section of this report will create unique resources and stable career opportunities that will make academic careers significantly more attractive. There is such high demand for advanced AI expertise that when universities have AI job openings they find themselves competing with a vigorous marketplace. This concerns not only the hiring of junior faculty, but the retention of senior tenured faculty. It is imperative to set up competitive academic career programs for doctorate-level researchers and early-career faculty, including multi-year fellowships and career initiation grants, in order to maintain the educational programs to train the next generation. In addition, strong programs that attract students to AI from early stages (high school and undergraduate) and promote AI careers should be a key ingredient for growing a workforce with advanced AI expertise.

5.2.3 ENGAGING UNDERREPRESENTED AND UNDERPRIVILEGED GROUPS

We recommend significant growth of participation by diverse, underrepresented and underprivileged groups, both to ensure access to the career opportunities afforded by this growing area and to significantly broaden the talent pool. This will require incorporation of best practices into funding processes and programs, across the board, in a comprehensive approach to diversity in AI that is consistently applied across educational and professional institutions and has extensive follow-up and effective redirection when needed:

- ▶ Incentivizing hiring and promotion processes to increase diversity, so that all well-qualified candidates have appropriate consideration and opportunities.
- ▶ Ensuring the diverse role models are present at all levels of AI instruction, from K-12 to undergraduate, graduate, and professional certificate training. The absence of women and members of other underrepresented groups in the education and training process results in perpetuating under-representation.
- ▶ Countering undesirable effects of common cultural practices, so that any disadvantages can be addressed up front. For example, women are more likely to take time off to care for children, so childcare programs and flexible job requirements should be commonplace.
- ▶ Providing access to college and career opportunities, including through bridge programs, so that any individual has the awareness, the knowledge, and the means required to pursue careers in AI. For example, a disproportionate number of minority students attend high schools that do not offer computer science courses, have limited access to computers at home, and lack mentors for careers in computer science in general, let alone in AI.
- ▶ Broadening the sources of student application pools, especially at the graduate level, so that students originate from a variety of backgrounds and institutions beyond traditional well-connected sources.
- ▶ Providing free AI coding training and opportunities to contribute to open-source AI software repositories. ROS is an excellent example of such a model and a major force multiplier in robotics: It is an international open-source community where volunteers can receive training and contribute code/software, which also serves as experience on resumes. GitHub code commits have similar value and are becoming the norm in software engineering training and hiring today.
- ▶ Developing and deploying targeted mentoring programs, cohort programs, and special interest groups that provide exposure to the AI field, individualized advice and coaching, and peer support.
- ▶ Explicitly codifying best practices in ethical behavior, conduct, and inclusiveness in academic, industry, and government organizations, so that inappropriate interactions, isolation, and implicit biases are eliminated from the school and the workplace.
- ▶ Confronting attrition on a continuous basis, so that the core underlying reasons for underrepresented groups leaving the AI field can be addressed quickly, systematically, and thoroughly.

A number of nonprofit organizations and grassroots efforts are successfully attracting K-12 students from underrepresented groups and underprivileged areas to STEM and computer science—particularly robotics—in large numbers. Also, a number of nonprofit organizations focus on mentoring college-age students and early-career professionals belonging to underrepresented groups to retain them in the computer science field. These could provide useful models for our recommendations regarding attracting a broader population of students to pursue long-term careers in AI. Using application domains such as AI for social good will serve as compelling attractors for women and other currently underrepresented communities of AI developers.

5.2.4 INCENTIVIZING EMERGING INTERDISCIPLINARY AI AREAS

We recommend the creation of courses and curricula to train students to pursue careers at the intersection of AI with different disciplines. This would:

- ▶ Train a larger and more diverse workforce with AI skills
- ▶ Enhance opportunities for AI for social good
- ▶ Serve AI needs in different sectors and application areas
- ▶ Accelerate the research priorities in this Roadmap
- ▶ Stimulate the generation of transformative ideas and approaches
- ▶ Promote research and applications of AI in all areas
- ▶ Enable multidisciplinary work that would advance our understanding of hard societal challenges for AI design and deployment involving AI ethics and safety

5.2.4.1 Interdisciplinary AI Studies

The kinds of AI innovations and challenges envisioned in this report will require a skilled workforce with a diversity of interdisciplinary backgrounds that only exists in very limited forms today.

True interdisciplinary work requires a deep understanding of other disciplines. Learning is more efficient and effective when new concepts in a related discipline are introduced in the context of what a student already knows. Interdisciplinary courses and curricula are hard to design, and best practices should be developed to guide their creation and implementation. It will also be important to provide incentives and articulate career paths to attract students to these programs.

We recommend promoting interdisciplinary AI studies, with opportunities to combine AI research with education. There are several aspects to this interdisciplinarity that we discuss in turn: the relevance of many sciences to the study of intelligence, the close ties of AI with other areas of computer science, and the need for discipline-specific and application-specific studies of AI.

Many AI researchers have traditionally worked at the intersection of AI and other fields of science. Understanding intelligence involves many other scientific disciplines, such as psychology, cognitive science, neuroscience, philosophy, evolutionary biology, sociology, anthropology, economics, decision science, linguistics, control theory, engineering, and mathematics among others. Significantly more interdisciplinary work and training programs are needed.

AI advances will come hand in hand with advances in computer science, and educational courses and programs can not only stimulate research but also create a workforce with diverse AI-relevant skills. AI advances will drive innovations in algorithms, hardware architectures, software engineering, distributed computing, data science, sensors and computing devices, user interfaces, privacy and security, databases, computer systems and networks, and theory, among others. Interdisciplinary programs and training opportunities will be important for the pursuit of the research agenda but also to support the comprehensive deployment across sectors. Furthermore, the technology will continuously change over time, and training will need to be designed so the workforce can easily adapt to such changes.

Important application areas of AI have been pursued in the past and have vibrant communities of both research and practice. These include AI and medicine, AI and education, AI and operations research, AI and manufacturing, AI and transportation, and AI and information science to name a few. Other areas are nascent and growing, such as AI and environmental sustainability, AI and law, AI for social sciences and humanities, AI for engineering, and AI for social good, among others. Academic programs that promote the understanding and application of AI in different sectors of importance to society need to be made more ubiquitous to meet increased workforce demand.

5.2.4.1.1 AI Ethics and Policy

The research priorities in this Roadmap highlight the importance of the area of AI ethics and policy, and the imperative of incorporating ethics and related responsibility principles as central elements in the design and operation of AI systems. Many of the challenges to be addressed are highly interdisciplinary, such as:

- ▶ **AI ethics** - concerned with responsible uses of AI, incorporating human values, respecting privacy, universal access to AI technologies, addressing AI systems bias particularly for marginalized groups, the role of AI in profiling and behavior prediction, as well as algorithmic fairness, accountability, and transparency. This research involves other disciplines such as philosophy, sociology, economics, science and technology studies, anthropology, and law.
- ▶ **AI privacy** - concerned with the collection and management of sensitive data about individuals and organizations, the use and protection of sensitive data in AI applications, and the use of AI for identification and tracking of people. This research involves other disciplines such as psychology and human factors, systems engineering, public policy, and law.
- ▶ **AI safety** - concerned with systems engineering to ensure reliable and safe operations, prevention of damage, resilience to failure, robustness, predictability, monitoring, and control. This research involves other disciplines such as industrial engineering and systems engineering.
- ▶ **AI policy** - concerned with legal aspects, regulation, and policy for AI deployments in different sectors and contexts. This research involves other disciplines such as public policy, government, and law.
- ▶ **AI trust** - concerned with high-stakes decision making, responsibility and accountability, individual's rights, and public perception and misconception of AI systems. This research involves other disciplines such as psychology, neuroscience, linguistics, and philosophy.

We recommend incentivizing the exploration of approaches that combine technology solutions, legal solutions, and policy solutions.

- ▶ Instrumentation of AI systems: Articulating the probes (how to instrument AI systems), indicators (what to observe), and metrics (how to measure) would enable a better quantitative characterization of key aspects of AI systems' behavior and level of autonomy to make decisions.
- ▶ Metrics are challenging to develop and with their effect, hard to anticipate. The domains in which AI systems are deployed are highly rich and dynamic, and a single metric can easily miss concerns that arise through complex effects.
- ▶ Carefully designed metrics can have a role, however, in informing a broader framework that should include legal, policy, and enforcement dimensions:
 - Legal aspects: The design of new laws that speak to concerns about agency, responsibility, and accountability, and the consequences of their violations.
 - Policy aspects: The design of new regulation and policy, as well as the characterization of desirable and permissible uses of AI in different sectors and application domains.
 - Enforcement aspects: Auditing and tracking compliance with existing policies and laws, as well as the detection of violations and undesirable outcomes.

We view that goal-driven instrumentation of AI systems, keeping in mind the larger legal, policy and enforcement goals, can contribute to the larger discussions around ethics and policy.

The mission-driven AI laboratories will provide a fertile ground for interdisciplinary research, for designing approaches that explore the complex space across these interrelated areas, and for investigating domain-specific requirements and solutions. More importantly, they will provide training and education opportunities at all levels for students and researchers in AI, public policy, legal studies, philosophy, and other areas, as well as practitioners to acquire the comprehensive background and understanding needed to tackle these challenging but crucial areas.

5.2.4.1.2 AI and the Future of Work

This is a key priority area of interdisciplinary research, concerned with disruption of the workforce, automation of jobs, emergence of new jobs, retraining, and other economic, social, and ethical impacts of the deployment of AI systems in the workplace. These challenges are at the intersection of AI with other disciplines such as economics, public policy, and education. Novel computing technologies often improve our lives, but they can also affect them in ways that are harmful or unjust. It is important to teach students how to think through the ethical and social implications of their work.

Great innovation and growth will result from making it easier for organizations to swiftly transform their existing processes into new ones that incorporate new technologies. AI is transforming existing jobs and creating new types of jobs, many of which will involve different forms of human-machine collaboration or assistance. A key research challenge is the characterization and management of organizational processes that combine AI systems and humans. These processes should articulate the different jobs in an organization, as well as the roles and responsibilities of people and AI systems in those jobs and processes. The characterization of organizational processes should support the identification of employee skill requirements, job reassignment, and retraining. Interdisciplinary research is needed to create proper characterizations of AI systems and human-machine systems within those processes. This will provide a useful framework to articulate the value of human roles in organizations and the limitations of automation on a task-by-task basis. It will also facilitate innovation in organizational processes and in the AI systems within them.

5.2.5 TRAINING HIGHLY SKILLED AI ENGINEERS AND TECHNICIANS

We recommend the training of significant numbers of AI engineers, as the demand is already high and will significantly increase over time. A significant portion of the AI workforce needed is for *AI engineers* who: 1) understand AI algorithms, data, and platforms; 2) set up and maintain AI infrastructure at scale; and 3) are able to continuously learn and adopt new technologies on the job. AI engineers combine basic expertise in AI with practical skills in software engineering, data systems, and high-end computing. These skills enable them to support the deployment of AI systems and the prototyping of AI innovations.

We also recommend training significant numbers of *AI technicians*. Companies already hire for such positions in large numbers, as these are people who can do data collection, data annotation, data cleaning, and other forms of data acquisition processes that are key for the development and improvement of AI systems. For example, AI technicians annotate failed interactions with conversational assistants, which generates data to train these AI systems to do better in the future. A workforce capability to supply skilled AI technicians could enable significant innovation across all sectors of the economy. For example, technicians could be trained to annotate sensor circuits, which could be used to train AI systems to support engineering designers. Workforce training programs and career opportunities should be created, particularly to handle complex data in diverse application areas and domains of interest.

We recommend that extensive adult education programs, distance education programs, and online courses be developed to fit personal circumstances and schedules of those interested in pursuing AI careers.

5.2.6 WORKFORCE RETRAINING

Workforce retraining programs to convert skilled technical workers in areas of dwindling demand into AI engineers and AI technicians will be greatly beneficial. This is particularly important for large portions of the workforce who currently have technical jobs and who will be able to repurpose their skills into AI. Other significant portions of the workforce that lack technical training will need special programs to enable them to access new jobs and opportunities in this emerging sector of the workforce.

A key contribution that AI can make in this arena is in the area of education and training. The research priorities in this Roadmap include significant aspects of AI for personalized education that broadens peoples' expertise in an area of work, retraining experts so they can repurpose their existing skills into new jobs, learning on the job to operate machines or to fit new organizational processes, and on-the-fly acquisition of skills. The fast pace of advances in AI will demand customizable, personalized, lifelong educational tools that can adapt to individual needs and incentives.

5.3 Core Programs for AI Research

The research priorities outlined in this Roadmap are driven by important societal drivers and require a broad range of programs. A wide variety of existing programs across agencies support AI research. However, the rate of progress in those research priorities will be slow if they continue at current funding levels and time horizons. In addition, many competitive research proposals on important areas of AI are not selected due to limited available funding. We recommend that funding for existing AI programs be increased, with particular emphasis for programs that support the research priorities in this Roadmap such as:

- ▶ Programs that support basic AI research: These programs support a single investigator (or two to three investigator collaborations) that would include students. Of particular importance are programs to support early career development and postdoctoral training.
- ▶ Programs that support application-driven AI research: These are programs that motivate AI research through the lens of practical problems.
- ▶ Interdisciplinary AI programs: These programs encourage researchers from different disciplines to pursue new approaches that combine AI, computing, engineering, social sciences, arts, and humanities. These programs should emphasize the diversity of ideas and participation, and should be for at least five years in order to reach their full potential.
- ▶ Programs representing private-public partnerships for AI: These programs would pursue unique opportunities facilitated by industry resources that can boost research in key areas of AI, such as computing and cloud resources, data resources, and other infrastructure.
- ▶ Integration of AI research and education: These programs would support the involvement of undergraduates and high school students in research projects. Programs should also be in place to incentivize their pursuit of advanced degrees.
- ▶ Outreach, diversity, and inclusion programs for AI: These programs would promote outreach and inclusion for underrepresented groups, and for institutions with limited capability in AI research and education. This would involve creating cohort groups with strong mentoring and longitudinal follow-up and building on ongoing work by nonprofit organizations and other grassroots efforts to engage broader segments of the population in AI topics.
- ▶ Education and curriculum development programs for AI: These programs would support the development of curriculum guidelines and educational materials, training opportunities for teachers and educators, online education programs for professionals, and incentives for careers in AI education.