

McKinsey Global Institute

# Applying artificial intelligence for social good

November 28, 2018 | Discussion Paper

By [Michael Chui](#), [Martin Harrysson](#), [James Manyika](#), [Roger Roberts](#), Rita Chung, Pieter Nel, and Ashley van Heteren

AI is not a silver bullet, but it could help tackle some of the world's most challenging social problems.

---

## DOWNLOADS

[↓ Discussion Paper \(PDF-3MB\)](#)

**A**rtificial intelligence (AI) has the potential to help tackle some of the world's most challenging social problems. To analyze potential applications for social good, we compiled a library of about 160 AI social-impact use cases. They suggest that existing capabilities could contribute to tackling cases across all 17 of the UN's sustainable-development goals, potentially helping hundreds of millions of people in both advanced and emerging countries.

Real-life examples of AI are already being applied in about one-third of these use cases, albeit in relatively small tests. They range from diagnosing cancer to helping blind people navigate their surroundings, identifying victims of online sexual exploitation, and aiding disaster-relief efforts (such as the flooding that followed Hurricane Harvey in 2017). AI is

only part of a much broader tool kit of measures that can be used to tackle societal issues, however. For now, issues such as data accessibility and shortages of AI talent constrain its application for social good.

This article is a condensed version of our discussion paper, [Notes from the AI frontier: Applying AI for social good](#) (PDF–3MB). It looks at domains of social good where AI could be applied, and the most pertinent types of AI capabilities, as well as the bottlenecks and risks that must be overcome and mitigated if AI is to scale up and realize its full potential for social impact. The article is divided into five sections:

1. [Mapping AI use cases to domains of social good](#)
2. [AI capabilities that can be used for social good](#)
3. [Overcoming bottlenecks, especially around data and talent](#)
4. [Risks to be managed](#)
5. [Scaling up the use of AI for social good](#)

# 1

## Mapping AI use cases to domains of social good

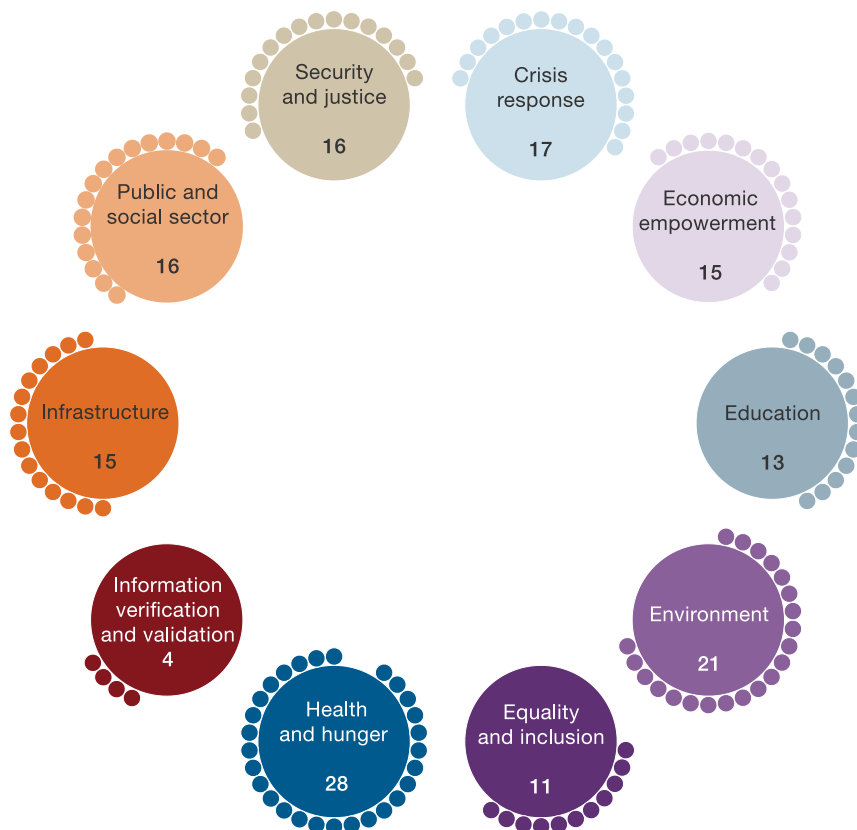
For the purposes of this research, we defined AI as deep learning. We grouped use cases into ten social-impact domains based on taxonomies in use among social-sector organizations, such as the AI for Good Foundation and the World Bank. Each use case highlights a type of meaningful problem that can be solved by one or more AI capability. The cost of human suffering, and the value of alleviating it, are impossible to gauge and compare. Nonetheless, employing usage frequency as a proxy, we measure the potential impact of different AI capabilities.

For about one-third of the use cases in our library, we identified an actual AI deployment (Exhibit 1). Since many of these solutions are small test cases to determine feasibility, their functionality and scope of deployment often suggest that additional potential could be captured. For three-quarters of our use cases, we have seen solutions deployed that use some level of advanced analytics; most of these use cases, although not all, would further benefit from the use of [AI techniques](#). Our library is not exhaustive and continues to evolve, along with the capabilities of AI.

#### Exhibit 1

Artificial intelligence (AI) has broad potential across a range of social domains.

AI use cases per domain, number



Note: Our library of about 160 use cases is not comprehensive and will continue to evolve. This listing of the number of cases per domain should thus not be read as exhaustive.

McKinsey&Company | Source: McKinsey Global Institute analysis

## Crisis response

These are specific crisis-related challenges, such as responses to natural and human-made disasters in search and rescue missions, as well as the outbreak of disease.

Examples include using AI on satellite data to map and predict the progression of wildfires and thereby optimize the response of firefighters. Drones with AI capabilities can also be used to find missing persons in wilderness areas.

## Economic empowerment

With an emphasis on currently vulnerable populations, these domains involve opening access to economic resources and opportunities, including jobs, the development of skills, and market information. For example, AI can be used to detect plant damage early through low-altitude sensors, including smartphones and drones, to improve yields for small farms.

## Educational challenges

These include maximizing student achievement and improving teachers' productivity. For example, adaptive-learning technology could base recommended content to students on past success and engagement with the material.

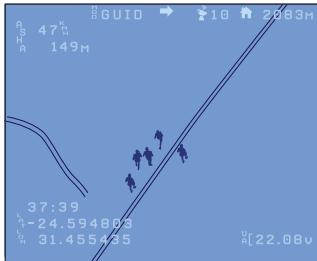
## Environmental challenges

Sustaining biodiversity and combating the depletion of natural resources, pollution, and climate change are challenges in this domain. (See Exhibit 2 for an illustration on how AI can be used to catch wildlife poachers.) The [Rainforest Connection](#), a Bay Area nonprofit, uses AI tools such as Google's TensorFlow in conservancy efforts across the world. Its platform can detect illegal logging in vulnerable forest areas by analyzing audio-sensor data.

Exhibit 2

Artificial intelligence (AI) can be deployed to catch wildlife poachers.

Six steps from offline training of AI model to online detection



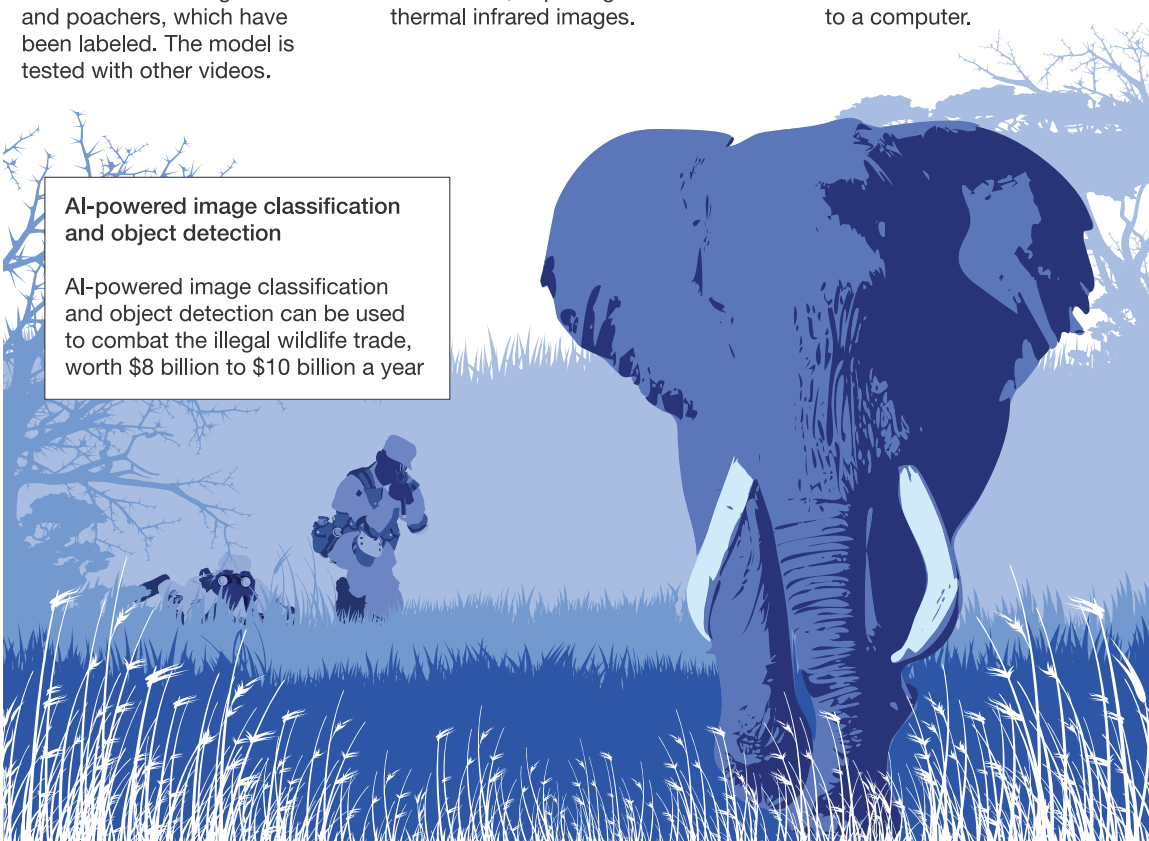
**1. Offline training**  
A neural network is trained on 70 videos containing animals and poachers, which have been labeled. The model is tested with other videos.



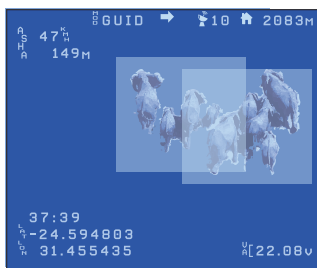
**2. Drone deployment**  
Drones are flown over wildlife sanctuaries, capturing thermal infrared images.



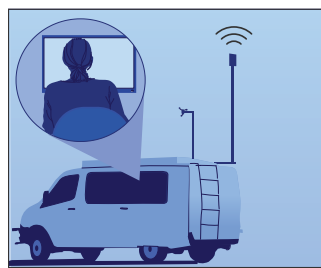
**3. User interface**  
Video and still images are transmitted via radio waves to a computer.



**AI-powered image classification and object detection**  
AI-powered image classification and object detection can be used to combat the illegal wildlife trade, worth \$8 billion to \$10 billion a year



**4. Preprocessing**  
The infrared images may need to be converted to “white-hot” format, where warm objects are lighter against a dark background.



**5. Detection**  
The video is processed in batches and sent to the cloud for analysis. Each image is treated as an input into the neural network.



**6. Output**  
The neural network outputs annotations that are overlaid on top of the original image. This enables identification of the poachers' whereabouts.

**Source:** Elizabeth Bondi et al., *SPOT poachers in action: Augmenting conservation drones with automatic detection in near real time*, 32nd AAAI Conference on Artificial Intelligence, April 27, 2018; McKinsey Global Institute analysis

McKinsey&Company

---

## Equality and inclusion

Addressing challenges to equality, inclusion, and self-determination (such as reducing or eliminating bias based on race, sexual orientation, religion, citizenship, and disabilities) are issues in this domain. One use case, based on work by Affectiva, which was spun out of the MIT Media Lab, and Autism Glass, a Stanford research project, involves using AI to automate the recognition of emotions and to provide social cues to help individuals along the autism spectrum interact in social environments.

## Health and hunger

This domain addresses health and hunger challenges, including early-stage diagnosis and optimized food distribution. Researchers at the University of Heidelberg and Stanford University have created a disease-detection AI system—using the visual diagnosis of natural images, such as images of skin lesions to determine if they are cancerous—that [outperformed professional dermatologists](#). AI-enabled wearable devices can already detect people with potential early signs of diabetes with 85 percent accuracy by analyzing [heart-rate sensor data](#). These devices, if sufficiently affordable, could help more than 400 million people around the world afflicted by the disease.

## Information verification and validation

This domain concerns the challenge of facilitating the provision, validation, and recommendation of helpful, valuable, and reliable information to all. It focuses on filtering or counteracting misleading and distorted content, including false and polarizing information disseminated through the relatively new channels of the internet and social

media. Such content can have severely negative consequences, including the manipulation of election results or even [mob killings, in India](#) and [Mexico](#), triggered by the dissemination of false news via messaging applications. Use cases in this domain include actively presenting opposing views to ideologically isolated pockets in social media.

## Infrastructure management

This domain includes infrastructure challenges that could promote the public good in the categories of energy, water and waste management, transportation, real estate, and urban planning. For example, traffic-light networks can be optimized using real-time traffic camera data and [Internet of Things \(IoT\)](#) sensors to maximize vehicle throughput. AI can also be used to schedule predictive maintenance of public transportation systems, such as trains and public infrastructure (including bridges), to identify potentially malfunctioning components.

## Public and social-sector management

Initiatives related to efficiency and the effective management of public- and social-sector entities, including strong institutions, transparency, and financial management, are included in this domain. For example, AI can be used to identify tax fraud using alternative data such as browsing data, retail data, or payments history.

## Security and justice

This domain involves challenges in society such as preventing crime and other physical dangers, as well as tracking criminals and mitigating bias in police forces. It focuses on security, policing, and criminal-justice issues as a unique category, rather than as part of public-sector management. An example is using AI and data from IoT devices to create solutions that help firefighters determine safe paths through burning buildings.

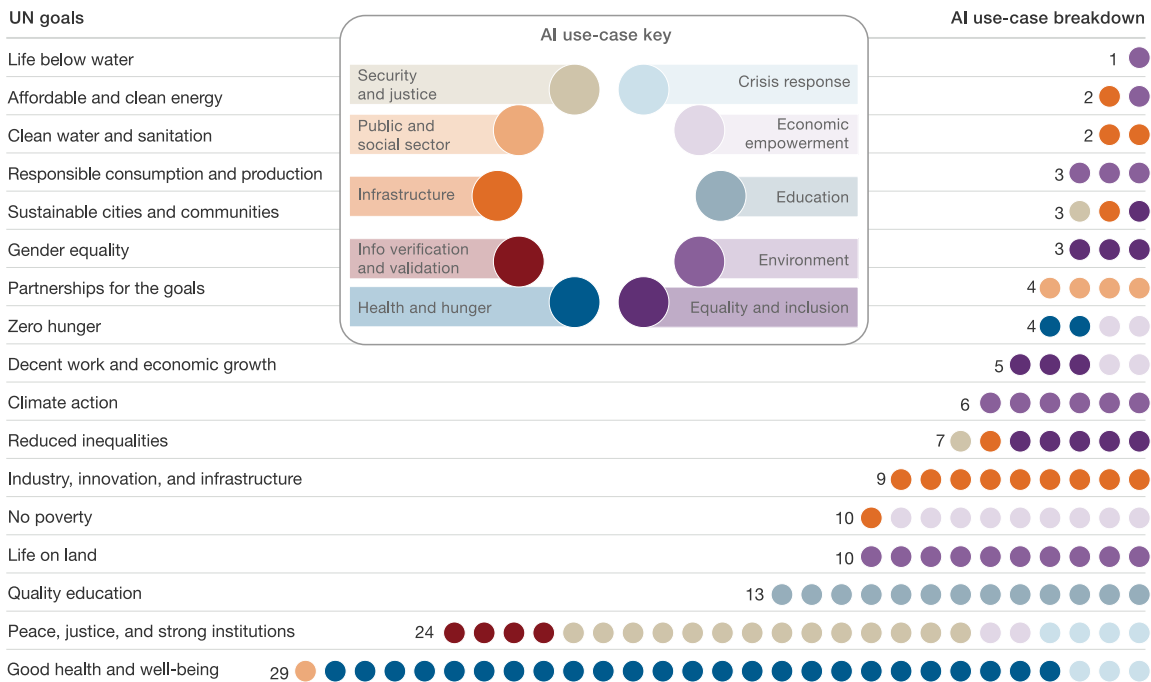
# Our use-case domains cover all 17 of the UN's Sustainable Development Goals

The United Nations' [Sustainable Development Goals](#) (SDGs) are among the best-known and most frequently cited societal challenges, and our use cases map to all 17 of the goals, supporting some aspect of each one (Exhibit 3). Our use-case library does not rest on the taxonomy of the SDGs, because their goals, unlike ours, are not directly related to AI usage; about 20 cases in our library do not map to the SDGs at all. The chart should not be read as a comprehensive evaluation of AI's potential for each SDG; if an SDG has a low number of cases, that reflects our library rather than AI's applicability to that SDG.

Exhibit 3

Many artificial-intelligence (AI) use cases support the most frequently cited societal challenges.

Where AI use cases fall within the UN's sustainable-development goals,<sup>1</sup> number



Note: This chart reflects the number and distribution of use cases and should not be read as a comprehensive evaluation of AI's potential for each sustainable-development goal (SDG); if an SDG has a low number of cases, that is a reflection of our library rather than of AI's applicability to that SDG. The chart also does not reflect all use cases in the library, more than 20 of which do not map to any SDG. These mainly focus on effective management in the public and social sectors, or belong to the issue types of disaster response and search and rescue in the crisis-response domain.  
<sup>1</sup>Chart is a partial list of use cases, as 21 of the 156 identified use cases do not target any of the UN's sustainable-development goals.

McKinsey&Company | Source: McKinsey Global Institute analysis



# 2

## AI capabilities that can be used for social good

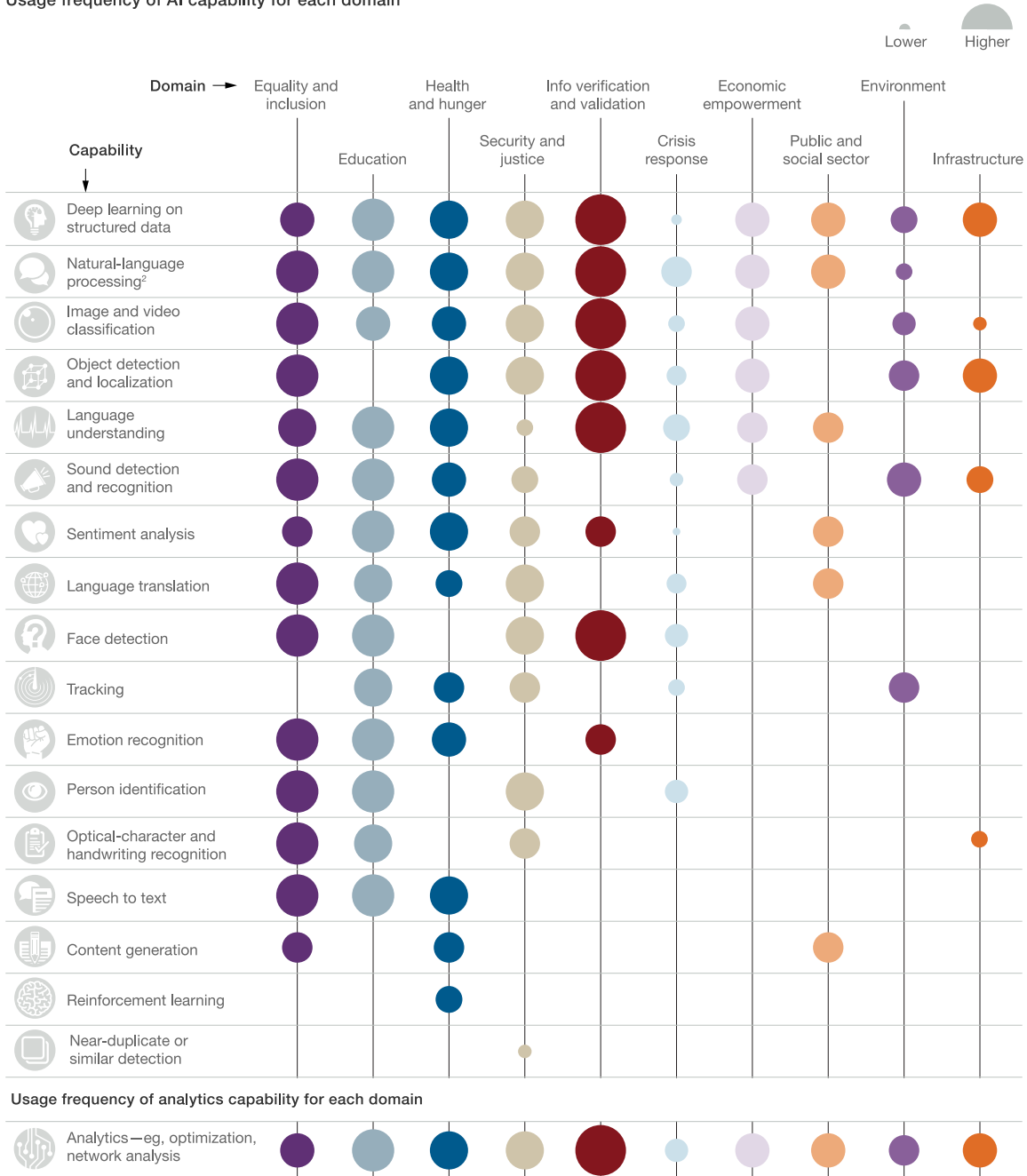
We identified 18 AI capabilities that could be used to benefit society. Fourteen of them fall into three major categories: computer vision, natural-language processing, and speech and audio processing. The remaining four, which we treated as stand-alone capabilities, include three AI capabilities: reinforcement learning, content generation, and structured deep learning. We also included a category for analytics techniques.

When we subsequently mapped these capabilities to domains (aggregating use cases) in a heat map, we found some clear patterns (Exhibit 4).

Exhibit 4

Artificial-intelligence (AI) capabilities mapped to their potential uses in ten domains where they may be of societal benefit.

Usage frequency of AI capability for each domain<sup>1</sup>



Note: Our library of about 160 use cases with societal impact is evolving and this heat map should not be read as a comprehensive gauge of the potential application of AI or analytics capabilities. Usage frequency estimates the number of times that models trained using AI would be used in a year to predict an outcome.

<sup>1</sup>Log base 10 scale. Deployment frequency capped at once per hour per year to prevent skewing; capping affected only a small number of use cases.

<sup>2</sup>Excluding sentiment analysis, speech to text, language understanding, and translation.

McKinsey&Company | Source: McKinsey Global Institute analysis

# Image classification and object detection are powerful computer-vision capabilities

Within computer vision, the specific capabilities of image classification and object detection stand out for their potential applications for social good. These capabilities are often used together—for example, when drones need computer vision to navigate a complex forest environment for search-and-rescue purposes. In this case, image classification may be used to distinguish normal ground cover from footpaths, thereby guiding the drone's directional navigation, while object detection helps circumvent obstacles such as trees.

Some of these use cases consist of tasks a human being could potentially accomplish on an individual level, but the required number of instances is so large that it exceeds human capacity (for example, finding flooded or unusable roads across a large area after a hurricane). In other cases, an AI system can be more accurate than humans, often by processing more information (for example, the early identification of plant diseases to prevent infection of the entire crop).

Computer-vision capabilities such as the identification of people, face detection, and emotion recognition are relevant only in select domains and use cases, including crisis response, security, equality, and education, but where they are relevant, their impact is great. In these use cases, the common theme is the need to identify individuals, most easily accomplished through the analysis of images. An example of such a use case would be taking advantage of face detection on surveillance footage to detect the presence of intruders in a specific area. (Face detection applications detect the presence of people in an image or video frame and should not be confused with facial recognition, which is used to identify individuals by their features.)

## Natural-language processing

Some aspects of natural-language processing, including sentiment analysis, language translation, and language understanding, also stand out as applicable to a wide range of domains and use cases. Natural-language processing is most useful in domains where information is commonly stored in unstructured textual form, such as incident reports, health records, newspaper articles, and SMS messages.

As with methods based on computer vision, in some cases a human can probably perform a task with greater accuracy than a trained [machine-learning](#) model can. Nonetheless, the speed of “good enough” automated systems can enable meaningful scale efficiencies—for example, providing automated answers to questions that citizens may ask through email. In other cases, especially those that require processing and analyzing vast amounts of information quickly, AI models could outperform humans. An illustrative example could include monitoring the outbreak of disease by analyzing tweets sent in multiple local languages.

Some capabilities, or combination of capabilities, can give the target population opportunities that would not otherwise exist, especially for use cases that involve understanding the natural environment through the interpretation of vision, sound, and speech. An example is the use of AI to help educate children who are on the autism spectrum. Although professional therapists have proved effective in creating behavioral-learning plans for children with autism spectrum disorder (ASD), [waitlists for therapy](#) are long. AI tools, primarily using emotion recognition and face detection, can increase access to such educational opportunities by providing cues to help children identify and ultimately learn facial expressions among their family members and friends.

## Structured deep learning also may have social-benefit applications

A third category of AI capabilities with social-good applications is structured [deep learning](#) to analyze traditional tabular data sets. It can help solve problems ranging from tax fraud (using tax-return data) to finding otherwise hard to discover patterns of insights in electronic health records.

Structured deep learning (SDL) has been gaining momentum in the commercial sector in recent years. We expect to see that trend spill over into solutions for social-good use cases, particularly given the abundance of tabular data in the public and social sectors. By automating aspects of basic feature engineering, SDL solutions reduce the need either for domain expertise or an innate understanding of the data and which aspects of the data are important.

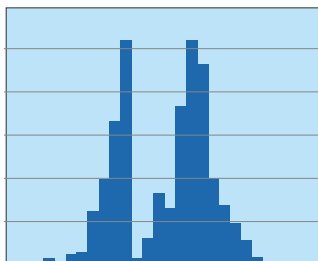
## **Advanced analytics can be a more time- and cost-effective solution than AI for some use cases**

Some of the use cases in our library are better suited to traditional analytics techniques, which are easier to create, than to AI. Moreover, for certain tasks, other analytical techniques can be more suitable than deep learning. For example, in cases where there is a premium on explainability, decision tree-based models can often be more easily understood by humans. In Flint, Michigan, machine learning (sometimes referred to as AI, although for this research we defined AI more narrowly as deep learning) is being used to predict houses that may still have lead water pipes (Exhibit 5).

Exhibit 5

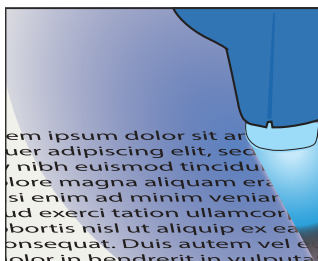
Artificial intelligence (AI) is being used to help detect lead pipes in Flint, Michigan.

Six steps from offline training of AI model to online detection



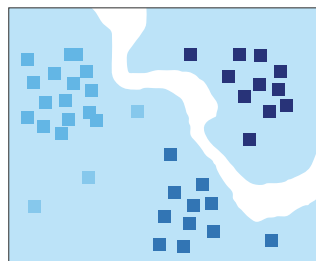
**1. Data collection**

City of Flint gave team data set describing each of the 55,893 parcels in city. It includes attributes of each home, such as owner, address, value, and building characteristics.



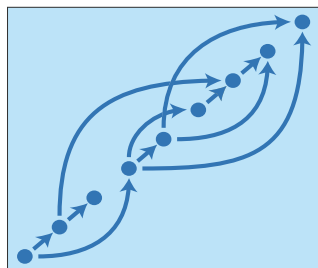
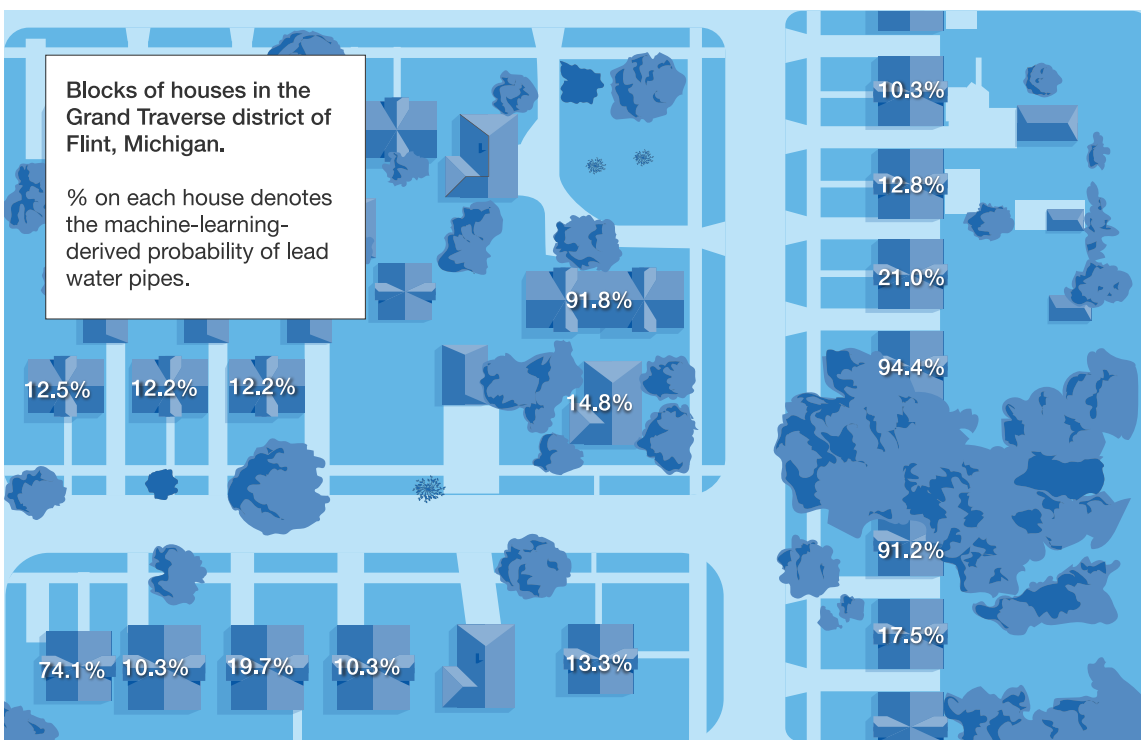
**2. Digitization**

Digitization of city service-line records using OCR on data, such as hand-annotated maps with markings for each parcel and incomplete records of materials for each home's service line.



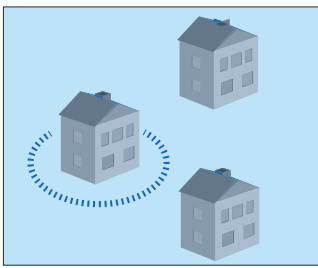
**3. Model building**

Machine-learning model predicts probability that the portion of a service line attached to a particular home is hazardous given known features, such as home value.



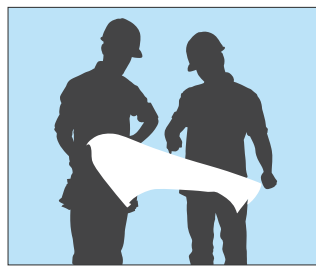
**4. Fine-tuning**

A Bayesian model is used to fine-tune predictions to account for relationships between parcels



**5. Detection**

Algorithm determines which homes to observe next, by efficiently allocating resources.



**6. Determination**

Determination of final list of homes to be visited by the replacement crew, based on

for relationships between parcels and precincts.

the probability that homes have lead pipes.

Source: Jacob Abernethy et al., *Active Remediation: The search for lead pipes in Flint, Michigan*, Cornell University Library, August 17, 2018; McKinsey Global Institute analysis

McKinsey&Company

# 3

## Overcoming bottlenecks, especially for data and talent

While the social impact of AI is potentially very large, certain bottlenecks must be overcome if even some of that potential is to be realized. In all, we identified 18 potential bottlenecks through interviews with social-domain experts and with AI researchers and practitioners. We grouped these bottlenecks in four categories of importance.

The most significant bottlenecks are data accessibility, a shortage of talent to develop AI solutions, and “last-mile” implementation challenges (Exhibit 6).

### Exhibit 6

Eighteen bottlenecks could limit artificial intelligence’s (AI’s) benefit to society.

#### Four categories of limitations to AI use

Critical barriers for most domains	Critical barriers for select cases <sup>1</sup>	Contextual challenges	Potential bottlenecks
<ul style="list-style-type: none"> <li>● Data accessibility</li> <li>● Data quality</li> <li>● High-level AI-expertise availability</li> <li>● High-level AI-expertise accessibility</li> <li>● Regulatory limitations</li> <li>● Organizational-deployment efficiency</li> </ul>	<ul style="list-style-type: none"> <li>● Data volume</li> <li>● Data labeling</li> <li>● AI-practitioner talent availability</li> <li>● AI-practitioner talent accessibility</li> <li>● Access to computing capacity</li> </ul>	<ul style="list-style-type: none"> <li>● Data availability</li> <li>● Data integration</li> <li>● Access to technology</li> <li>● Privacy concerns</li> <li>● Organizational receptiveness</li> </ul>	<ul style="list-style-type: none"> <li>● Access to software libraries and other tools</li> <li>● Organizations able to scale AI deployment</li> </ul>

Note: This list of bottlenecks was derived from interviews with social-sector experts and AI researchers and tested against our use cases. <sup>1</sup>Bottlenecks that are critical for some domains as a whole or for individual use cases within those domains.

McKinsey&Company | Source: McKinsey Global Institute analysis

## Data needed for social-impact uses may not be easily accessible

Data accessibility remains a significant challenge. Resolving it will require a willingness, by both private- and public-sector organizations, to make data available. Much of the data essential or useful for social-good applications are in private hands or in public institutions that might not be willing to share their data. These data owners include telecommunications and satellite companies; social-media platforms; financial institutions (for details such as credit histories); hospitals, doctors, and other health providers (medical information); and governments (including tax information for private individuals). Social entrepreneurs and nongovernmental organizations (NGOs) may have difficulty accessing these data sets because of regulations on data use, privacy concerns, and bureaucratic inertia. The data may also have business value and could be commercially available for purchase. Given the challenges of distinguishing between social use and commercial use, the price may be too high for NGOs and others wanting to deploy the data for societal benefits.

## The expert AI talent needed to develop and train AI models is in short supply

Just over half of the use cases in our library can leverage solutions created by people with less AI experience. The remaining use cases are more complex as a result of a combination of factors, which vary with the specific case. These need high-level AI expertise—people who may have PhDs or considerable experience with the technologies. Such people are in short supply.

For the first use cases requiring less AI expertise, the needed solution builders are data scientists or software developers with AI experience but not necessarily high-level expertise. Most of these use cases are less complex models that rely on single modes of data input.



The complexity of problems increases significantly when use cases require several AI capabilities to work together cohesively, as well as multiple different data-type inputs. Progress in developing solutions for these cases will thus require high-level talent, for which demand far outstrips supply and competition is fierce.

## **‘Last-mile’ implementation challenges are also a significant bottleneck for AI deployment for social good**

Even when high-level AI expertise is not required, NGOs and other social-sector organizations can face technical problems, over time, deploying and sustaining AI models that require continued access to some level of AI-related skills. The talent required could range from engineers who can maintain or improve the models to data scientists who can extract meaningful output from them. Handoffs fail when providers of solutions implement them and then disappear without ensuring that a sustainable plan is in place.

Organizations may also have difficulty interpreting the results of an AI model. Even if a model achieves a desired level of accuracy on test data, new or unanticipated failure cases often appear in real-life scenarios. An understanding of how the solution works may require a data scientist or “translator.” Without one, the NGO or other implementing organization may trust the model’s results too much: most AI models cannot perform accurately all the time, and many are described as “brittle” (that is, they fail when their inputs stray in specific ways from the data sets on which they were trained).

# 4

## **Risks to be managed**

AI tools and techniques can be misused by authorities and others who have access to them, so principles for their use must be established. AI solutions can also unintentionally harm the very people they are supposed to help.

An analysis of our use-case library found that four main categories of risk are particularly relevant when AI solutions are leveraged for social good: bias and fairness, privacy, safe use and security, and “explainability” (the ability to identify the feature or data set that leads to a particular decision or prediction).

Bias in AI may perpetuate and aggravate existing prejudices and social inequalities, affecting already-vulnerable populations and amplifying existing cultural prejudices. Bias of this kind may come about through problematic historical data, including unrepresentative or inaccurate sample sizes. For example, AI-based risk scoring for criminal-justice purposes may be trained on historical criminal data that include biases (among other things, African Americans may be unfairly labeled as high risk). As a result, AI risk scores would perpetuate this bias. Some AI applications already show large disparities in accuracy depending on the data used to train algorithms; for example, [examination of facial-analysis software](#) shows an error rate of 0.8 percent for light-skinned men; for dark-skinned women, the error rate is 34.7 percent.

One key source of bias can be poor data quality—for example, when data on past employment records are used to identify future candidates. An AI-powered recruiting tool used by one tech company was abandoned recently after several years of trials. It appeared to show systematic bias against women, which resulted from patterns in training data from years of hiring history. To counteract such biases, skilled and diverse data-science teams should take into account potential issues in the training data or sample intelligently from them.

## Breaching the privacy of personal information could cause harm

Privacy concerns concerning sensitive personal data are already rife for AI. The ability to assuage these concerns could help speed public acceptance of its widespread use by profit-making and nonprofit organizations alike. The risk is that financial, tax, health, and similar records could become accessible through porous AI systems to people without a legitimate need to access them. That would cause embarrassment and, potentially, harm.

## **Safe use and security are essential for societal good uses of AI**

Ensuring that AI applications are used safely and responsibly is an essential prerequisite for their widespread deployment for societal aims. Seeking to further social good with dangerous technologies would contradict the core mission and could also spark a backlash, given the potentially large number of people involved. For technologies that could affect life and well-being, it will be important to have safety mechanisms in place, including compliance with existing laws and regulations. For example, if AI misdiagnoses patients in hospitals that do not have a safety mechanism in place—particularly if these systems are directly connected to treatment processes—the outcomes could be catastrophic. The framework for accountability and liability for harm done by AI is still evolving.

## **Decisions made by complex AI models will need to become more readily explainable**

Explaining in human terms the results from large, complex AI models remains one of the [key challenges](#) to acceptance by users and regulatory authorities. Opening the AI “black box” to show how decisions are made, as well as which factors, features, and data sets are decisive and which are not, will be important for the social use of AI. That will be especially true for stakeholders such as NGOs, which will require a basic level of transparency and will probably want to give clear explanations of the decisions they

make. Explainability is especially important for use cases relating to decision making about individuals and, in particular, for cases related to justice and criminal identification, since an accused person must be able to appeal a decision in a meaningful way.

## Mitigating risks

Effective mitigation strategies typically involve “human in the loop” interventions: humans are involved in the decision or analysis loop to validate models and double-check results from AI solutions. Such interventions may call for cross-functional teams, including domain experts, engineers, product managers, user-experience researchers, legal professionals, and others, to flag and assess possible unintended consequences.

Human analysis of the data used to train models may be able to identify issues such as bias and lack of representation. Fairness and security “red teams” could carry out solution tests, and in some cases third parties could be brought in to test solutions by using an adversarial approach. To mitigate this kind of bias, university researchers have demonstrated methods such as sampling the data with an understanding of their inherent bias and creating synthetic data sets based on known statistics.

Guardrails to prevent users from blindly trusting AI can be put in place. In medicine, for example, misdiagnoses can be devastating to patients. The problems include false-positive results that cause distress; wrong or unnecessary treatments or surgeries; or, even worse, false negatives, so that patients do not get the correct diagnosis until a disease has reached the terminal stage.

Technology may find some solutions to these challenges, including explainability. For example, nascent approaches to the transparency of models include [local-interpretable-model-agnostic](#) (LIME) explanations, which attempt to identify those parts of input data a trained model relies on most to make predictions.

# Scaling up the use of AI for social good

As with any technology deployment for social good, the scaling up and successful application of AI will depend on the willingness of a large group of stakeholders—including collectors and generators of data, as well as governments and NGOs—to engage. These are still the early days of AI's deployment for social good, and considerable progress will be needed before the vast potential becomes a reality. Public- and private-sector players all have a role to play.

## Improving data accessibility for social-impact cases

A wide range of stakeholders owns, controls, collects, or generates the data that could be deployed for AI solutions. Governments are among the most significant collectors of information, which can include tax, health, and education data. Massive volumes of data are also collected by private companies—including satellite operators, telecommunications firms, utilities, and technology companies that run digital platforms, as well as social-media sites and search operations. These data sets may contain highly confidential personal information that cannot be shared without being anonymized. But private operators may also commercialize their data sets, which may therefore be unavailable for pro-bono social-good cases.

Overcoming this accessibility challenge will probably require a global call to action to record data and make it more readily available for well-defined societal initiatives.

Data collectors and generators will need to be encouraged—and possibly mandated—to open access to subsets of their data when that could be in the clear public interest. This is already starting to happen in some areas. For example, many satellite data companies participate in the [International Charter on Space and Major Disasters](#), which commits

them to open access to satellite data during emergencies, such as the September 2018 tsunami in Indonesia and Hurricane Michael, which hit the US East Coast in October 2018.

Close collaboration between NGOs and data collectors and generators could also help facilitate this push to make data more accessible. Funding will be required from governments and foundations for initiatives to record and store data that could be used for social ends.

Even if the data are accessible, using them presents challenges. Continued investment will be needed to support high-quality data labeling. And multiple stakeholders will have to commit themselves to store data so that they can be accessed in a coordinated way and to use the same data-recording standards where possible to ensure seamless interoperability.

Issues of data quality and of potential bias and fairness will also have to be addressed if the data are to be deployed usefully. Transparency will be a key for bias and fairness. A deep understanding of the data, their provenance, and their characteristics must be captured, so that others using the data set understand the potential flaws.

All this is likely to require collaboration among companies, governments, and NGOs to set up regular data forums, in each industry, to work on the availability and accessibility of data and on connectivity issues. Ideally, these stakeholders would set global industry standards and collaborate closely on use cases to ensure that implementation becomes feasible.

## **Overcoming AI talent shortages is essential for implementing AI-based solutions for social impact**

The long-term solution to the talent challenges we have identified will be to recruit more students to major in computer science and specialize in AI. That could be spurred by significant increases in funding—both grants and scholarships—for tertiary education

and for PhDs in AI-related fields. Given the high salaries AI expertise commands today, the market may react with a surge in demand for such an education, although the advanced math skills needed could discourage many people.

Sustaining or even increasing current educational opportunities would be helpful. These opportunities include “AI residencies”—one-year training programs at corporate research labs—and shorter-term AI “boot camps” and academies for midcareer professionals. An advanced degree typically is not required for these programs, which can train participants in the practice of AI research without requiring them to spend years in a PhD program.

Given the shortage of experienced AI professionals in the social sector, companies with AI talent could play a major role in focusing more effort on AI solutions that have a social impact. For example, they could encourage employees to volunteer and support or coach noncommercial organizations that want to adopt, deploy, and sustain high-impact AI solutions. Companies and universities with AI talent could also allocate some of their research capacity to new social-benefit AI capabilities or solutions that cannot otherwise attract people with the requisite skills.

Overcoming the shortage of talent that can manage AI implementations will probably require governments and educational providers to work with companies and social-sector organizations to develop more free or low-cost online training courses. Foundations could provide funding for such initiatives.

Task forces of tech and business translators from governments, corporations, and social organizations, as well as freelancers, could be established to help teach NGOs about AI through relatable case studies. Beyond coaching, these task forces could help NGOs scope potential projects, support deployment, and plan sustainable road maps.

From the modest library of use cases that we have begun to compile, we can already see tremendous potential for using AI to address the world’s most important challenges. While that potential is impressive, turning it into reality on the scale it deserves will require focus, collaboration, goodwill, funding, and a determination among many

stakeholders to work for the benefit of society. We are only just setting out on this journey. Reaching the destination will be a step-by-step process of confronting barriers and obstacles. We can see the moon, but getting there will require more work and a solid conviction that the goal is worth all the effort—for the sake of everyone.

---

## ABOUT THE AUTHOR(S)

**Michael Chui** is a partner and **James Manyika** is chairman and a director of the McKinsey Global Institute. **Martin Harrysson** and **Roger Roberts** are partners in McKinsey's Silicon Valley office, where **Rita Chung** is a consultant. **Pieter Nel** is a specialist in the New York office; **Ashley van Heteren** is an expert associate principal in the Amsterdam office.

[Talk to us](#)

---

## EXPLORE A CAREER WITH US

[Search Openings](#)