The Spectrum of Artificial Intelligence
Companion to the FPF AI Infographic
Contents

The Spectrum of Artificial Intelligence
Companion to the FPF AI Infographic

I. EXECUTIVE SUMMARY 4
   A. Symbolic AI 4
   B. Machine Learning 5
   C. Risks and Benefits of AI 5

II. INTRODUCTION 6

III. FOUNDATION DISCIPLINES 7
    Philosophy 7
    Ethics 7
    Logic 8
    Mathematics 8
    Physics 8

IV. MODERN COMPONENTS 8
    A. Data 8
    B. Statistics 9
    C. Design 9
    Security 9
    Hardware 9

V. ARTIFICIAL INTELLIGENCE (NON-ML) – OVERVIEW 10
    A. Rules Based AI 10
    B. Symbolic AI 11
       i. Search 11
       ii. Planning and Scheduling 11
       iii. Expert Systems 12
    C. Computer Sensing 12
    D. Robotics 13
    E. Knowledge Engineering 13
    F. Natural Language Processing 14
    G. Risks and Benefits – Rules-based AI 14

VI. ARTIFICIAL INTELLIGENCE (MACHINE LEARNING) 15
    A. Reinforcement Learning 15
    B. Neural Networks 16
    C. Deep Learning 17
    D. Generative Adversarial Networks 17
    E. Risks and Benefits – Machine Learning 17

VII. CONCLUSION 19

AUTHORED BY

Brenda Leong
Senior Counsel & Director of Artificial Intelligence and Ethics

Dr. Sara R. Jordan
Senior Researcher, Artificial Intelligence and Ethics

COVER ART

Artist: Dr. Lydia Kostopoulos
Title: Luncheon of the Tech Enabled Boating Party
Year: 2020
Artificial Intelligence (AI) is the computerized ability to perform tasks commonly associated with human intelligence, including reasoning, discovering patterns and meaning, generalizing, applying knowledge across spheres of application, and learning from experience. The growth of AI-based systems in recent years has garnered much attention, particularly in the sphere of Machine Learning. A subset of AI, Machine Learning (ML) systems "learn" from the success or accuracy of their outputs, and can change their processing over time, with minimal human intervention. But there are non-ML types of AI that, alone or in combination, lie behind the real-world applications in common use. General AI — a human-level computational system — does not yet exist. But Narrow AI exists in many fields and applications where computerized systems greatly enhance human output or outperform humans at defined tasks. This chart explains the main types of AI, their relationships to each other, and provides specific examples of how they are currently appear in our day-to-day lives. It also demonstrates how AI exists within the timeline of human knowledge and development.
THE SPECTRUM OF ARTIFICIAL INTELLIGENCE

I. Executive Summary

This paper is a companion piece to the FPF Spectrum of Artificial Intelligence (AI) Infographic (ffp.org/blog/the-spectrum-of-artificial-intelligence-an-infographic-tool), to expand the information included in that educational resource, and describe how the graphic can be used as an aid in developing legislation or other regulatory guidance impacting AI-based systems. We identify specific use cases for various approaches and technologies and highlight the differing algorithmic architecture and data demands present varying risks and benefits. We discuss the spectrum of algorithmic approaches and demonstrate how design factors, data use, and model training processes should be considered for specific regulatory approaches.

Recent calls for regulation of AI-based systems exist within the complex landscape of contemporary AI programs with a variety of approaches from agricultural crop growth detectors to automated book recommendations. Some approaches focus on a need to cover “algorithms” generally while other initiatives are narrower, suggesting regulation of AI specifically as used in automated vehicles or medical devices. This paper aims to assist the development of any of these approaches by making clear some of the terms, relationships, and functions of AI, and how they might be impacted by differing restrictions. Any regulatory efforts should be made with the best understanding of how different algorithms, use cases, applications, and ultimately individuals, will be affected. Despite the media focus on Machine Learning (ML), there are many other types of AI preceding and operating alongside ML, all of which have different attributes and aspects, and which will need differently formulated regulatory controls or guidance.

Many forms of artificial intelligence are the clear analogues to human thought processes or human physical actions which can be integrated within systems to help humans in a variety of ways that are precise or consistently, and with more information than an individual could. But, of course, AI systems can also process large amounts of data beyond what any human could do, to identify patterns or predict outcomes that are far beyond what conventional data science and statistical methods could accomplish.

A. Symbolic AI

Traditional Rules-Based AI leading to Symbolic AI (sometimes used synonymously) is the collection of all methods in artificial intelligence research that are based on human understandings (“commonsense”) of mathematics, logic, and coded programming. Symbolic AI systems represent the first significant steps towards designing machines to enable complex decisions or reason through complexity and uncertainty. There are several algorithmic designs that are generally considered to be examples of Symbolic AI, including Search, Planning and Scheduling, and Expert Systems.

When computers are programmed to find a specific pattern in a set of symbols and then to perform a designed action, we can say that the AI is engaged in a “search.” Planning and scheduling AI are what enable a computer to take into account multiple dimensions that require adjustment of strategies, such as playing a video game (e.g., gaming points) while avoiding traps (e.g., losing lives). Search AI and planning and scheduling AI play important roles in the boring and hidden parts of systems that provide many of the conveniences in modern life. For example, supply chain management, including airline cargo scheduling systems and “just in time” restocking models rely on these programs. Expert systems identify solutions by combing through and combining multiple types and layers of information, using the reasoning and logic common to a particular profession or specialty, such as medicine or engineering. An expert system can provide faster, more reliably accurate diagnoses based on personal health data, or design recommendations for pollution abatement given the environmental and industrial factors involved.

Building computers that can “see,” “listen,” “smell,” or “taste” as ways to evaluate their physical environment requires new approaches to computer sensing that generally rely on a combination including several other forms of AI as well. Computers can play an important role as a building block for creating augmented intelligence used in advanced and assistive robotics.

Teaching AI to reason using the same cognitive patterns as expert professionals involves teaching machines the bases of professional common knowledge laid out as a set of underlying rules for processing and establishing expectations, knitting together the wealth of documents, rules, and common knowledge that people rely on as “knowledge engineering” systems, they use rules and pattern recognition to sift through data such as tax codes, extract relevant patterns, and categories, and provide an expert’s guidance. For example, this analysis could formulate question and answer sets to guide an individual through their tax return preparation, following the appropriate steps for that individual’s finances.

Natural language processing (NLP) systems are some of the most common AI systems that people routinely encounter. Powering home-based assistants, and various devices or appliances; providing language translation tools, predictive typing, autocorrect, question and answer systems, and robotic speech, such as Siri, Alexa, or Google Assistant. They can read written texts, and comparing drafts for plagiarism, and even writing independent, creative work. NLP combines ML with other forms of AI in everyday systems and tools.

B. Machine Learning

The types of AI described above tell computers how to sift through information according to rules and processes crafted by humans, such as language or mathematics. Machine learning is different. Machine learning works because machines use an initial set of rules (programming) to identify connections and patterns which they then use to internally edit their instructions or build additional rules of their own. Computers using reinforcement and prior analyses to improve subsequent calculations or minimize loss of performance is what makes machine learning so powerful. Machine learning has advantages and traditionally difficult AI tasks, such as image recognition, and provides the ability to analyze constantly changing information flows for applications like social media content monitoring.

Neural networks, the building blocks of some types of machine learning, learn by identifying patterns within input data to make new, internal, rules about the relationships between the data and outputs. These systems allow computers to process highly complex information quickly, sometimes approaching human levels of association and “intuition.” These networks can be layered to process data through multiple programs sequentially and repeatedly for more sophisticated analysis. When they are layered, they may comprise a “deep learning” system. These are the systems trained to recognize objects, such as color values, edges, and commonly associated items so that the output value, such as probability that a specific image is a canoe and not a cat, can be provided to the user.

Most recently, there are two newer forms of machine learning enabling powerful systems to achieve major advances: generative adversarial networks (GANs) and reinforcement learning (RL). Reinforcement learning is a key step in designing AI to independently learn human-like, goal-oriented, tasks. Reinforcement learning is what powered the system that learned to master the game “Go.” The first ML system, which defeated the human world Go champion 4 games to 1, operated on traditional machine learning processes. The next program was designed using reinforcement learning and beat that original system, 100-0. Reinforcement learning systems will likely power the next generation of robotics, for purposes such as search and rescue missions in complex environmental situations or high-capability home care assistants.

GANs, the newest variation of machine learning AI being developed, are based on creating a pair of neural networks that learn by attempting to better each other: first, the “generator” of the pair creates an output (e.g., an image) based upon the initial human programming. The other network, the “discriminator,” has been programmed to what the correct output should be (e.g., what the image should look like). The discriminator evaluates the output, and critiques it. Initial outputs are likely to be extremely inaccurate. The discriminator’s feedback is then incorporated, the generator continues to churn out results, and the feedback loop continues until the generator produces data that the discriminator believes meets the quality expectations. These GANs type of learning is what drives “deep fakes” and some entertainment uses of AI and will likely inform or improve other systems in the future.

C. Risks and Benefits of AI

The future of AI ultimately lies in the goals and systems towards which humans direct it. Responsible uses should include two primary foundations for AI: to further advance human knowledge and to improve human lives. AI is key to the future of knowledge in many scientific disciplines and commercial technologies but carries accompanying risks that it will be applied unfairly, or designed unfairly, and that individuals and groups will be worse off in specific or personal ways. However, the potential benefits are powerfully significant, if sufficient effort is applied, and the fair and beneficial impacts for a greater social good.

AI systems operate across a broad spectrum of scale. Processes using these technologies can be designed to seek solutions to macro level problems like environmental challenges around undetected earthquakes, pollution control, and other natural disaster responses while they are also incorporated into personal level systems for greater access to educational, economic, and professional opportunities. If reinforcement is to be effective, it should focus on both technical details and the underlying values and rights that must be protected from adverse uses of AI, to ensure that AI is ultimately used to promote human dignity and welfare.
II. Introduction

This paper is a companion piece to the FPF Spectrum of Artificial Intelligence (AI) Infographic, to further explain the information included in this educational resource, and it should be used as an aide in developing any legislation around AI-based systems.

If one of the many calls to regulate AI were successful, what exactly would be regulated? In many ways, the answer to this question is “it depends.” Some approaches start from the premise that AI is a discrete type of software technology, like an operating system, and assume regulations should focus on preventing unfair impacts on things like credit scoring. Other approaches define AI as an entire system of hardware plus software, like a robotic arm or a personal home assistant, and should be regulated in ways based on safety concerns, similar to modern connected automobiles.

As we will discuss, the contemporary spectrum of AI is broad, and any call to regulate AI must align regulatory controls that are appropriate to the context, and in light of the specific harms of the systems being considered.

Artificial intelligence is a term with a long history. Meant to denote those systems which accomplish tasks otherwise understood to require human intelligence, AI is directly connected to the development of computer science but is based on a myriad of academic fields and disciplines, including philosophy, social science, physics, mathematics, logic, statistics, and ethics. As AI as it is designed and used today is made possible by the recent advent of unprecedentedly large datasets, increased computational power, advances in data science, machine learning, and strong AI vs weak AI models, it includes programming and system design based on a number of sub-categories, such as robotics, expert systems, scheduling and planning systems, natural language processing, neural networks, and machine learning. In many cases of consumer facing AI, multiple forms of AI are used together to accomplish the overall performance goal specified for the system. In addition to considerations of algorithmic design, data flows, and programming languages, AI systems are most robust for use in equitable and stable consumer uses when human designers also consider potential hardware, cybersecurity, and user-interface design.

Two arguments are most commonly behind calls for regulation. The first is that AI presents unique risks that are not present in prior technological advancements. Under this argument, AI regulations should introduce mechanisms to identify and mitigate potentially harmful outcomes, both tangible and intangible, as well as offer solutions to rectify situations where poorly managed or unforeseen risks caused damage to humans, the environment, or social good.

The second, narrower argument focuses on the unknown risks that may arise due to the innate design aspects of AI that appear to be out of the control of the programmer or user of the system. Under this argument, because some forms of AI are designed in a way that humans cannot fully comprehend, explain, or reproduce, we should use AI with regulatory-driven caution until we can be more confident of the reliability or accuracy of these systems. Both of these arguments assume there are unique aspects to AI-based systems that are not, or cannot, be addressed by existing legal protections, regulatory schemes, or traditional policy approaches.

Our purpose is not to assert if or how AI systems should be regulated, but rather to provide a general understanding of the variety of AI systems that may be behind various applications and industries, and what the impacts might be, and to demonstrate that any regulatory action should be taken thoughtfully and in response to the particular areas of concern. Blunt approaches that seek to include any and all systems using AI-based models are considerably more likely to have unforeseen secondary consequences. Toward that end, this paper outlines the spectrum of AI technology, from rules-based and symbolic AI to advanced, developing forms of neural networks, and seeks to put them in the context of other sciences and disciplines, as well as emphasize the importance of security, user interface, and other design factors. Additionally, we seek to make this understandable through providing specific use cases for the various types of AI and by showing how the different architecture and data demands present specific risks and benefits.

Across the spectrum, AI is a combination of various types of reasoning. Rules-based or Symbolic AI is the form of algorithmic design wherein humans draft a context for how program behaviors, and rules that apply, are to be informed by data. In machine learning, neurons and machine learning systems are able to infer algorithms that can then be informed by data. These algorithms can then be trained to reason probabilistically or approximately may be used in cases where a “good enough” answer helps augment human decision-making. Also like recipes, AI systems can range from simple to extremely complex, with some of the most advanced systems being combinations of multiple algorithms creatively combined.

III. Foundation Disciplines

AI and Machine Learning (ML) are relatively new features of the scientific landscape, but they are built upon a long history of philosophical and scientific developments, including areas such as philosophy, ethics, logic, mathematics, and physics. In more recent decades, scientific disciplines that inform AI include data analytics, statistical modeling, and cybersecurity and encryption. Furthermore, these systems cannot be evaluated without also including considerations about the hardware devices and networks, the underlying logics and principles, and user interface and user experience designs. All of these areas contribute to the questions, answers, and analyst needed to fully review present day AI. While some might seem remote, or tangential, in fact the underlying values and assumptions of the designers and users are key to understanding the contextual implications of any particular AI system.
IV. Modern Components

A. Data

"Data" is the organized record and presentation of information. Modern "big data" in the digital world is part of the trail created by people, businesses or other organizational entities as they interact with sensors, surveys, sites, and applications embedded in the array of internet or computer-based, or computer-connected, products, services, and features. Data in this context has measurable characteristics like volume, variability, accuracy, and value, and also can have organizing principles around protection, privacy, provenance, and proportionality. In discussions about AI and ML, personal data is often categorized according to the sector from which it is gathered, such as healthcare, financial, educational, or consumer data. Data is also described based on the technological associations, including sensor data, location data, or visualized data. Data can also be synthetic, imputed, or meta (data about data). Data is not just the invaluable input into many AI and ML systems, but also includes the information created by or describing the analysis inherent in those systems, and the outputs and recommendations the systems produce.

The techniques necessary for processing data for uses in AI and ML have spurred both theoretical and applied research in areas of data management, including the ethical processing of data (e.g., differential privacy) and the technical processing techniques like dimensionality reduction (e.g., principal component analysis). The demand for AI-ready data has also spawned an industry for data engineering, a new form of AI-adjacent special demand for AI-ready data has also spawned an industry for data engineering, a new form of AI-adjacent special

Cleaning data, identifying bias in data, adding noise to data, deidentifying data, standardizing machine-readable data, analyzing data, and "owning" data are just a few of the many issues surrounding the collection, flow, and use of all of this information in modern digital systems.

B. Statistics

A branch of applied mathematics often tailored into specific fields (e.g., statistical mechanics, biostatistics), statistics inform the way in which the inputs for AI systems are evaluated and how the outputs are interpreted. Statistical modeling is the design of a way to make sense of data through analysis. Without established statistical theory, we would lack many of the essential concepts to explain how AI works or to evaluate how accurately machines have identified patterns. Established tools and measures such as variance, correlation, confidence, probability, and hypothesis testing inform the algorithmic models in order to glean insights and inform evidence-based reasoning. This reasoning drives how systems combine data to test representations of reality against one another, and provides the tools necessary to evaluate the results.

C. Design

While not always the first area of consideration for AI systems, the physical design of AI systems includes considerations of aesthetic value, creativity in physical presentation, and the efficiency of human interface. Apple, for example, is particularly known for its focus on beauty and intuitive function as part of its innate design structure. "Design thinking" — a process for brainstorming collaboratively — includes 5 stages: Empathize, Define, Prototype and Test and is deployed widely in the software and applications development landscape in user interface design and user experience design. More than simple efficiency or functionality, however, design considerations for AI systems must also include the understanding that they are integral components of applications that affect the lives of individual people, communities, cultures, and the environment. Whether focusing on personal privacy, bias and fairness, civil rights impacts, or other social impacts, design choices are a fundamental consideration in any AI system.

B. Statistics

A branch of applied mathematics often tailored into specific fields (e.g., statistical mechanics, biostatistics), statistics inform the way in which the inputs for AI systems are evaluated and how the outputs are interpreted. Statistical modeling is the design of a way to make sense of data through analysis. Without established statistical theory, we would lack many of the essential concepts to explain how AI works or to evaluate how accurately machines have identified patterns. Established tools and measures such as variance, correlation, confidence, probability, and hypothesis testing inform the algorithmic models in order to glean insights and inform evidence-based reasoning. This reasoning drives how systems combine data to test representations of reality against one another, and provides the tools necessary to evaluate the results.

C. Design

While not always the first area of consideration for AI systems, the physical design of AI systems includes considerations of aesthetic value, creativity in physical presentation, and the efficiency of human interface. Apple, for example, is particularly known for its focus on beauty and intuitive function as part of its innate design structure. "Design thinking" — a process for brainstorming collaboratively — includes 5 stages: Empathize, Define, Prototype and Test and is deployed widely in the software and applications development landscape in user interface design and user experience design. More than simple efficiency or functionality, however, design considerations for AI systems must also include the understanding that they are integral components of applications that affect the lives of individual people, communities, cultures, and the environment. Whether focusing on personal privacy, bias and fairness, civil rights impacts, or other social impacts, design choices are a fundamental consideration in any AI system.
Artificial intelligence has been part of the programming landscape, built on the academic foundational disciplines described above, since at least the 1940s. The earliest forms of AI are still in use today, including within some of our most widely used systems. In fact, examples of rule-based AI, such as those which make up the backbone of navigation systems, are so commonplace we often do not think of them as artificial intelligence any longer.

The initial drive to build artificial intelligence grew from the human desire to automate work that is repetitive, tedious, dangerous, requires high levels of precision, or is simply impossible for an individual or group of humans given some combination of factors. Many of these tasks are dependent on heuristics — formal and informal rules — that “everyone” within that field or domain just “knows.” Some of the first examples of rules-based or heuristic AI emerged in fields like chemical analysis, infectious disease diagnostics, or oncology diagnostics.

What was relevant about these domains was the belief that the logical structure of the questions to be asked and the specific types of information necessary to answer those questions was both available and could be symbolically represented to a computer. The rules for reaching “good” decisions, and the expert knowledge of uncertainty or confidence in the utility of a specific piece of information, could be systematically organized and ranked to produce a reliable, consistent answer by a non-human program that was on par with, or superior to, that deduced or provided by the human expert.

Rules-based systems maneuver through data using an “inference engine” — a set of logical commands according to which a computer interprets information and relates that information to the set of possible outputs (e.g., a probability score) in order to answer the questions asked of it. Because of the logic and symbolic structure of inference engines driving rule-based systems, they can work both backwards and forwards.

Inference engines work by either “forward chaining” or “backward chaining” through the rules. In a forward chaining strategy, an inductive approach moving from individual facts to general theory/moves forward from data inputs to conclusive outputs by matching the data to the best rule to identify a match. In a backward chaining strategy, a deductive logical approach (from general theory to individual facts) moves backward from provided set of possible outputs, through the rules and data, to identify if a particular output can be supported as legitimate. When a known set of possible states or decisions is generally well characterized, such as in healthcare diagnostic systems, a backward chaining strategy is useful. Where there is a wider or uncertain range of possible outputs from a system such as selecting, combining, and providing the appropriate mix of medications, a forward chaining strategy is useful.

Artificial intelligence covers many combinations of hardware and software components, but at its core, means a specific context, application, use case, or circumstance. Thus, the greatest chess player, and the world Go champion, are now AI systems. These systems, while far outstripping human capabilities in one specific area, are fairly use-less for almost any other purpose. While some aspects of their programming and sub-routes might be reused to jump start other projects, it is not a system that can intrinsically “learn” a new skill (function or application) without human involvement to edit and reapply that code, provide new data and training, and so on.

Across the spectrum, AI is a combination of various types of reasoning. Rules-based or Symbolic AI is that form of algorithmic design wherein humans draft a complete program of logical, connected commands for a computer to follow.41 Newer AI advances, particularly in machine learning systems based on neural networks, are able to power computers that carry out the programmer’s initial commands but then adapt their operations based on what the system can glean from patterns in the data. These systems evaluate their results and then connect those outcomes back into the code in order to improve the success of succeeding iterations of the program.42

Recently, what the media and many discussions around AI are most commonly referring to is actually machine learning, as if those two terms were exactly interchange-able. In fact, however machine learning as a specific subset of AI, Explaining the specifics of machine learning and the distinction between it and other types of AI is one of the primary goals of this paper and infographic. All ML is a form of AI, but not all AI is ML. This is one of the key takeaways to inform potential policymakers in their consideration of AI systems. We already should correctly identify what type of systems are being used interchangeably, and how those operate, so that standards, guidance, and restrictions can be targeted appropriately.

A. Rules-Based AI

The terms Symbolic AI and Rules-based AI are many times used interchangeably. Systems that automate processes in an imitation of human reasoning, such as through the use of logical statements or identification of patterns, are a type of Symbolic AI. Computers can be designed to detect patterns such as sequences of letters, numbers or other symbols, and then to reproduce an arrangement of those symbols so that humans can evaluate them. There are various versions of this type of system, which we are using as the umbrella term for the following three sub-categories:

i. Search

Search algorithms are used by many people in contemporary life. Since the beginning of the age of internet search engines, “search” has come to be synonymous with online browser search capabilities, but that is not the entirety of what this category includes. Search algorithms are used in many systems that are designed to find the best path to a goal within a specific set of possible solutions. Search algorithms are fundamental to the success of succeeding iterations of the program. These systems evaluate their results and then connect initial commands but then adapt their operations based on what the system can glean from patterns in the data. These systems, while far outstripping human capabilities in one specific area, are fairly use-less for almost any other purpose. While some aspects of their programming and sub-routes might be reused to jump start other projects, it is not a system that can intrinsically “learn” a new skill (function or application) without human involvement to edit and reapply that code, provide new data and training, and so on.

B. Symbolic AI

For more complex sets of variables that include constraints, such as finding a route to a city airport without using highway, or by a particular car method, an AI will normally be programmed by combining Search algorithms with Planning and Scheduling programming.

Planning algorithms are advanced Search algorithms that treat problem solving more like the way a human would. These planning algorithms can account for a variety of situational constraints, such as incomplete information, time constraints, and non-redundancy of resources. These types of models are used to design work plans, strategic design, and logistics planning for tasks like helping the Hubble Space Telescope arrange its scans of...
iii. Expert Systems

Some of the first Symbolic AI systems helped experienced practitioners in various professional domains to make, or teach others to make, complicated decisions, thus giving them the name “expert systems.” Expert systems identify possible solutions by combing through and combining multiple types and layers of information. Expert systems sift through information using the reasoning and logic common to a particular profession or specialty, such as medicine or engineering.

The key components of expert systems are an extensive, detailed “knowledge base,” the “inference engine,” and the available working memory. The knowledge base is the network of rules, logical statements, and domain specific knowledge and reasoning that defines the expertise of humans. The knowledge base is an engineered product constructed to organize the rules, norms, protocols, and standards of the specialty, including ranking and organizing them in sequence of the information needed. This organization is done using a system of symbols comprehensible to both humans and computers such as natural language dictionaries for natural language processing, or established rules for maneuvering around a defined environment, such as a chess board.

Thus, expert Systems augment the decision-making capacity of professionals in a specific topical domain by systematizing their most expert levels of knowledge into a chess board.

Perhaps the most intriguing use of AI and machine learning today is the ability to design and train computer programs to “sense” and then evaluate the physical environment around them. Computers can be designed with a range of sensors that simulate seeing, hearing, smelling, even tasting, and thus can be trained to collect, process, and respond to active or passive stimuli extending well beyond that of human sensory perception in both scope and accuracy. These sensors might perceive and measure light (including infrared spectra), touch, distance (range), temperature, acceleration, and speed. Although early programmers assumed that having computers interpret their environments visually would be one of the easiest tasks, creating the algorithms essential for computer vision has proven to be an extraordinarily difficult challenge that now routinely combines aspects of rules-based and symbolic AI together with machine learning.

Furthermore, computers are being trained to “hear,” “smell,” or even “feel.” Differentiating and classifying sounds is taught to computers in much the same way that “seeing” is. A similar set of algorithms, combined with neural networks (discussed under Machine Learning, below) are designed to analyze and classify sounds. Teaching computers to “smell” means to classify the molecules detected, and makes use of older forms of AI, combinations of AI and machine learning, and more recently, a newly designed form of neural networks called “graph neural networks.”

Beyond the physical senses, there is research designed to train computers to both perceive and demonstrate emotive behavior from or on par with humans. However, attempting to design models which can reliably identify, mimic, or initiate emotive behavior remains in the early stages, is currently imprecise and inconsistent, and even where accurate, triggers some of the most challenging ethical and social questions.

V. Artificial Intelligence – Overview (continued)

E. Knowledge Engineering

Knowledge Engineering is a field of AI oriented to build systems that emulate the judgment and behavior of a human expert by codifying knowledge as rules and relationships between data. These systems represent knowledge as directed acrylic graphs which are able to express complex calculations and logical eligibility rules. The graphs can be easily queried and the results reasoned to automatically produce a calculation or decision result. When reasoning using a Knowledge Engineering system, a backward chaining starts from the goal and works backward to determine what facts must be asserted so that the goal can be achieved.

As an example, Knowledge Engineering powers the programs designed to provide individual and business users with the ability to comply with the ever-changing taxation rules and regulations. Because there are so many and they change so frequently to greater and lesser degrees, it would be nearly impossible for an individual to effectively identify, extract, and reconcile the new rules against the old ones. To manage this at scale for people and businesses generally, an ensemble of rule-based algorithmic approaches are adopted: Natural Language Processing is used to review current laws and extract pertinent information; graph algorithms, such as networked analysis, can show the relationship of new rules to previous instances of the rule and also reflect the impact of other applicable rules. The information extracted can be further processed into ontologies (representations of abstract concepts) that establish the terms to encode for use in subsequent applications, such as those which forecast taxable income streams and associated revenue for states and the federal government. And thus the whole is created to guide a user through a detailed but highly individualized process based on their own information, against the background of the most current rules.
However, as useful as symbolic systems are, they do have limitations. For example, search and planning-driven navigation systems may be unable to optimize for the fastest route in a way that accounts for all user preferences, such as external safety considerations or for roads with a history of potholes. And like any AI system, adverse consequences may arise from flaws in the model components or the data, whether from errors in the knowledge base or inaccuracies within the inference engine for knowledge engineering. These systems’ highest value is most often as they augment humans already carrying out these processes. Expert systems facilitate more accurate and faster diagnoses and some treatment processes but have not replaced the need for human doctors. While domain knowledge can be carefully engineered, the process of maintaining and expanding a knowledge base to account for new information introduces complexity, overlap, and the potential for errors within control structures. And even expert systems, in the complex arena of medicine and diagnostics, keep a knowledge base current, and the potential that patients will not receive the latest, best, or most effective treatments.

There are also considerations like risks to physical safety when general planning algorithms are not appropriately tuned by human designers, such as those used to direct industrial robots or other automated industrial processes. Likewise, planning algorithms such as those which might make up the logistics strategy for delivering pandemic vaccines or other crisis response materials, may not provide optimal plans if they are not written to sufficiently account for all the challenges involved. Such as remote areas, prioritizing population cohorts for banded distribution, and storage and expiration — some of which may be difficult to define in new contexts or environments where information is not available. As with all automated processes, the quality of the outputs will always be driven by the quality of the design and the availability and accuracy of data inputs, rooted entirely in the systems for which information is not available. As with all automated processes, the quality of the outputs will always be driven by the quality of the design and the availability and accuracy of data inputs, rooted entirely in the systems for which information is not available.

The primary types of machine learning training are supervised and unsupervised learning, along with variations like semi-supervised, self-supervised, multi-task, and transfer learning. They may be defined as each has its own model of choices that lead to rewards. Nevertheless, transfer learning is increasingly common as companies look to maximize their return on investment. Particularly given the environment is itself a set of rules, as in the case of a car’s driverless algorithms for making better recommendations. The systems identify patterns in a dataset, edit inferential rules based on those patterns and connections, and then judge the fitness of that model’s outputs against the requirements set by the human programmers. When we say that a machine “learns,” it means the program iterates the process enough times to perform better and better, to make the model outputs more accurate against a previously set standard for success.

Machine learning has a number of sub-categories as well. ML systems can identify patterns in data we already know something about, such as when it classifies or predicts based on similarities of existing preferences, and can also derive conclusions from data we know nothing about yet, such as when it clusters items based on unseen connections or even generates new data. While there are general methods for training machines and designing machine learning systems, many programs are ultimately unique to context, designed for a specific problem. These programs predict, classify, categorize, mine, and learn from the data. Even machine learning program that start out initially become unique because each has “learned” through its own data-driven experiences, influenced by the initial weight and importance values assigned by the programmers, and then modified over time with real data.

The advancements of algorithmic systems are the various forms of Machine Learning (ML). As mentioned, the attention and public focus on these systems as they have begun to pervade everyday devices and systems have been so prevalent in recent years that many people, ML and AI are treated as one and the same. With the advent of mobile devices such as smartphones and tablets, machine learning has been implemented to provide everything from weather apps to video and book recommendations, personalized healthcare and fitness platforms, and programs to accelerate and support employment and educational needs. The average consumer may enjoy using these applications but will likely struggle to describe what machine learning is, how it works, or where it contributes to the applications found in their phones, televisions, or even appliances.
Neural Networks (NN) are a type of machine learning inspired by the human brain. They differ from other forms of machine learning in that they are non-linear by design.46 A NN consists of a web of interconnected entities known as nodes; each node carries out a simple computation, similarly to the neurons in the human brain. Neural Networks are a collection of the algorithms used in Machine Learning for data modeling operating within this graph of nodes. Data passes through several layers of interconnected nodes, as each node classifies the characteristics and information of the previous layer before passing the results on to other nodes in subsequent layer. The layers of networks pass the data through hierarchies of various concepts, which, like other machine learning models, allows them to learn through evaluating their own errors.

The first, or input layer, is where data is received, such as data uploaded to a cloud service or live sensor data, and is then subject to a mathematical function that manipulates the data for further use over multiple iterations. The input layer to a learning system is not simply data ingestion but is an active path of calculation. The input layer leads to (at least one) hidden layer, and finally to an output layer. Each layer contains one or more nodes. By increasing the number or complexity of the hidden layers, you increase the computational and problem-solving abilities of the model.47 Each node classifies the characteristics and information of the previous layer before passing the results on to other nodes in subsequent layer. The layers of networks pass the data through hierarchies of various concepts, which, like other machine learning models, allows them to learn through evaluating their own errors.

For this reason, CNNs are used in object detection and classification systems. CNNs can detect and differentiate objects within an image, such as road signs, cars. They commonly operate by finding the “edges” of different objects, and then comparing the arrangement of edges, with other images they’ve previously identified.48 When this doesn’t work as accurately as desired, undulating sand dunes may be mistaken for a reclining human body, for example, objects or faces in a 11 or many fashion, depending on the design of the system. In other words, if the system is trying to find all the traffic lights in the image, then every item is simply “yes, traffic light,” or “no, not traffic light.” Other systems may want to identify as many items in the image as possible, without knowing in advance what they might be.

In contrast, Recurrent Neural Networks (RNNs) learn from data where timing and sequencing are important features, to answer questions around forecasting — predicting changes in air quality, the next word in a sentence, or stock market performance estimates, based on numerous highly variable factors. RNNs begin with data from the input layer passed through the hidden layers to the output layers, however, they include loops within the hidden layers that mimic a type of short-term memory of what has already been processed. RNNs can also map a single input to multiple forms of output. This means that RNNs are well suited for problems such as translation (e.g., what the image should look like). The discriminator considers the output, and critiques it — essentially determining whether it is real, or correct. Early in the cycle, the initial outputs are likely to be far off from what is desired. The discriminator’s feedback is then incorporated into the program and the generator continues to churn out results, and the feedback loop continues, until the generator produces images that the discriminator believes meets the quality expectations. In the case of an image, this is the generator “fooling” it into thinking a generated output is real. GANs are very new and their capabilities are still being explored, but they are the systems used, so far, to produce “deep fakes,” contemporary works of art, music, or writing in the style of long dead masters, or to create entirely unique compositions by the AI system.

The input layer to a learning system is not simply data ingestion but is an active path of calculation. The input layer leads to (at least one) hidden layer, and finally to an output layer. Each layer contains one or more nodes. By increasing the number or complexity of the hidden layers, you increase the computational and problem-solving abilities of the model.47 Each node classifies the characteristics and information of the previous layer before passing the results on to other nodes in subsequent layer. The layers of networks pass the data through hierarchies of various concepts, which, like other machine learning models, allows them to learn through evaluating their own errors.

The first, or input layer, is where data is received, such as data uploaded to a cloud service or live sensor data, and is then subject to a mathematical function that manipulates the data for further use over multiple iterations. The input layer to a learning system is not simply data ingestion but is an active path of calculation. The input layer leads to (at least one) hidden layer, and finally to an output layer. Each layer contains one or more nodes. By increasing the number or complexity of the hidden layers, you increase the computational and problem-solving abilities of the model.47 Each node classifies the characteristics and information of the previous layer before passing the results on to other nodes in subsequent layer. The layers of networks pass the data through hierarchies of various concepts, which, like other machine learning models, allows them to learn through evaluating their own errors.

Reinforcement learning algorithms are what drive some of the notable forms of machine learning, such as when machines learn to play games, operate a flight simulator, or power a robot vacuum through a new environment.45 And because reinforcement learning does not require an existing large set of data on which to train or operate, it can avoid some of the technical and ethical challenges involved with creating and maintaining such datasets.
The ethics of public service delivery.

VII. Conclusion

AI is a field of science that encompasses technical, so-
cial, and policy considerations. As with any technology, there is no “neutral” system—the choices made of what to automate; how to determine the type, style, and
features included; sourcing the data; and applying the sys-
tem in specific contexts are all design choices that
carry implications for equity and harm. Automating hu-
manship carries with it inherent risks of automating human errors and shortcomings.

However, AI systems can also add specific and extend-
ed value to many fields, providing efficiencies in cost and
time, as well as accuracy and reliability. Mobile ser-

c

En
do
tes

1 Tenha Taleriya, Obari Shah, Nivedita Patel; Hitesh Yagnik, Manan Sinha, Priya Bhardwaj. “What is artificial intelligence in agriculture for
optimisation of irrigation and application of pesticides and

Driving Systems: A vision for safety”. Available at: https://

3 Ruhanka, G., & Thirumurthy, H. (2021). “Artificial Intelligence and
Machine Learning (AI/ML) as Software as a Medical Device
Planning”. Available at: https://www.fda.gov/medical-devices/software-medical-device-ai-machine-learning-software-medical-device

4 An important cultural note must be pointed out before examination of these foundational disciplines. Just as artificial intelligence does not belong to one culture or computer programming to one language, philosophy and mathematics do not exist only in cultures where CNNs are applied to image models, they can be
highly successful, reliably recognizing not only that a dog
is not a cat but that a Labrador retriever is not a Pekingese.
However, if poorly trained, they will distinguish wolves from dogs simply on the basis of snow in the picture, or mis-
categorize men and women of particular races, or with
specific health conditions. When moving beyond image

5 Jordan, Sara R., and Philip W. Gray. The ethics of public
administration. The challenges of global governance. Baylor
University Press, 2017; Jordan, Sara R. “The innovation imperative: An

6 Macintyre, Alasdair. A Short History of Ethics: a history of moral
philosophy from the Homeric age to the 20th century. Routledge,

7 Schuhler, Erin Anne, Sara R. Jordan, and Betty Barry. “Regulating


9 Apt, Krzysztof R. “Logic Programming.” Handbook of Theoretical
Computer Science. Volume B: Formal Models and Semantics (B)

10 Desrochers, Marc Peter; A Also Faisal and Ching Soon Ong.

11 Hogg, Tad, and Bernando A. Huberman. “Artificial intelligence and

12 Inside Big Data. 2020. “The 6 types of data everybody should know
to avoid confusion”. Available at: https://insidebigdata.com/2020/02/25/
the-6-types-of-data-everybody-should-know-to-avoid-confusion/


14 Khan, M Al-Alli-din, Muhammed Faisal Udun, and Nawam Gupta.
“Seven V’s of Big Data: Understanding big data to extract
value”. Available at: http://www.asee.org/documents/zones/adams2.0_090617_v9a_tag.pdf


18 FUTURE OF PRIVACY FORUM | AUGUST 2021

THE SPECTRUM OF ARTIFICIAL INTELLIGENCE | COMPANION TO THE FPF AFFI DOCUMENT
The Future of Privacy Forum (FPF) is a catalyst for privacy leadership and scholarship, advancing responsible data practices in support of emerging technologies. FPF is based in Washington, DC, and includes an advisory board comprising leading figures from industry, academia, law, and advocacy groups.