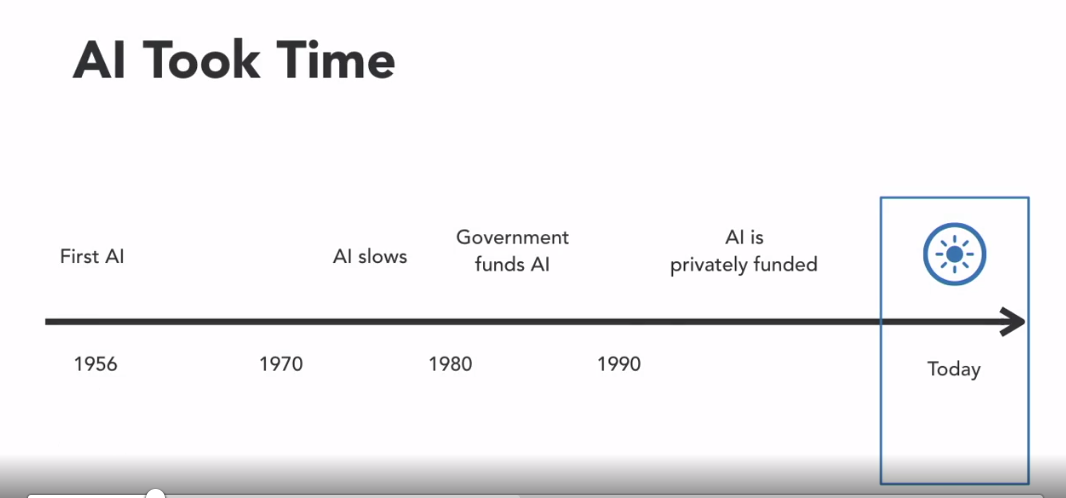
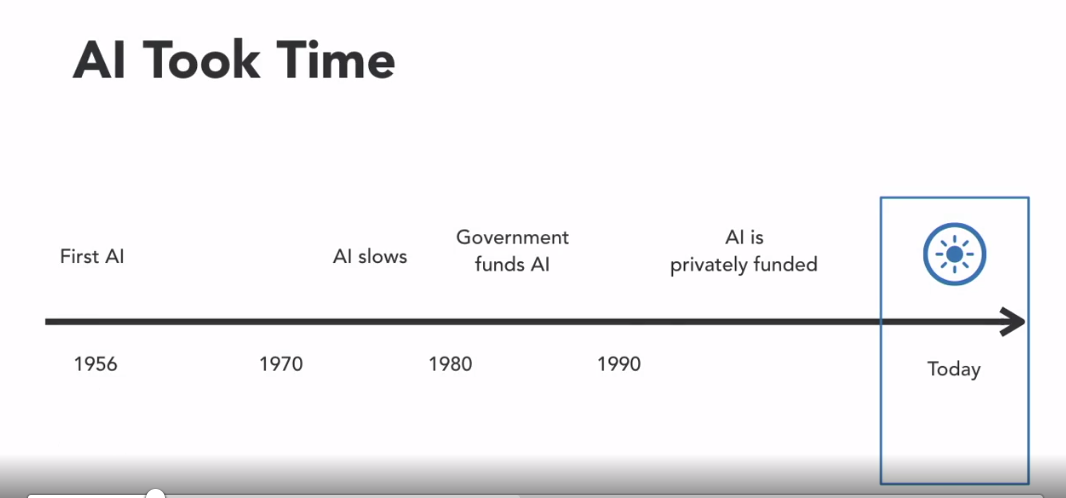
# **Learning XAI: Explainable Artificial Intelligence**

### **Explainable AI: Expanding the frontiers of artificial intelligence**

Selecting transcript lines in this section will navigate to timestamp in the video

- [Aki] Welcome, I am very excited that you decided to join me today for this course on Explainable AI. Have you ever asked questions like which is smarter, humans or AI? What might be the next big thing after AI? Is it possibles for humans and computers to get along better, or what in the world is Explainable AI? Then stay tuned, we're going to dig into the fascinating world of AI. Their strengths and weaknesses, the ways that humans can better work together with them, and, of course, what Explainable AI is. Although AI can be a powerful tool, one of its biggest drawbacks is that it cannot tell you why it made a certain decision or recommendation. It is essentially a black box. XAI is a type of AI that will also give us insight into the reasons why an outcome was recommended. In this course, we'll talk about what it is, how it works, why you might use it, how you could benefit from it, and what you should do about it. My name is Aki Ohashi, and I am a Director of Business Development for the Palo Alto Research Center, or PARC, the Xerox Research Center that invented technology such as the personal computer, laser printer, ethernet, and the graphical user interface. Today, PARC continues to conduct research in some of the world's most advanced technologies, including printed electronics, optoelectronics, AI and machine learning, Blockchain, as well as Explainable AI. At PARC, I oversee business development for all of our clients in Asia-Pacific, and Australia and New Zealand. With some of the largest companies in the world, many of these companies are already deeply involved in AI technology, and some have started their journey into XAI. We've got a lot to cover, so let's get going.

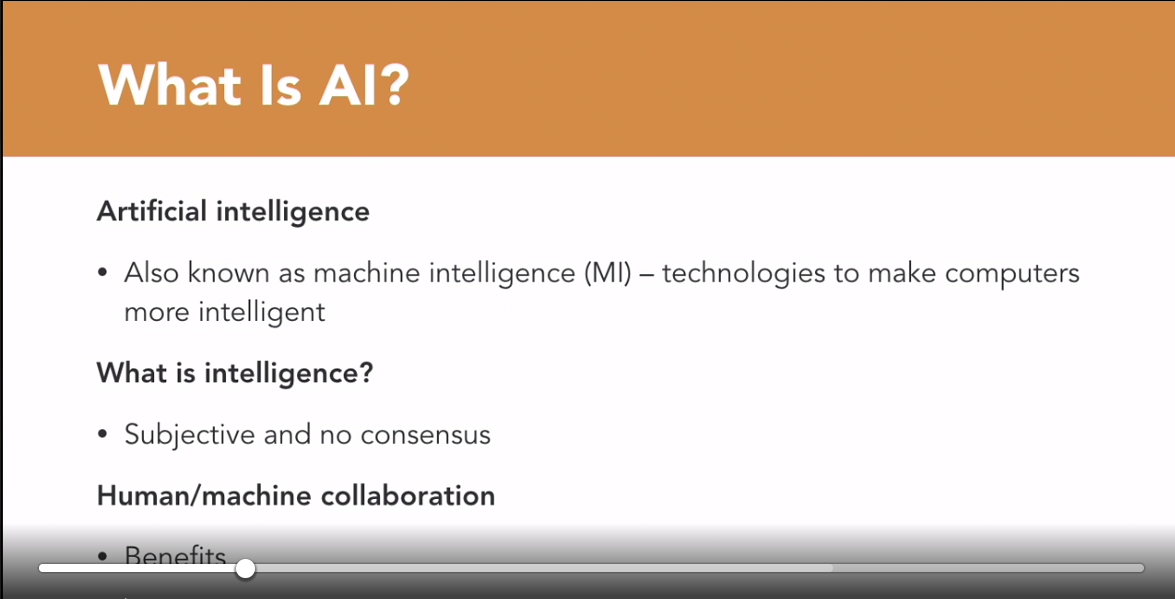


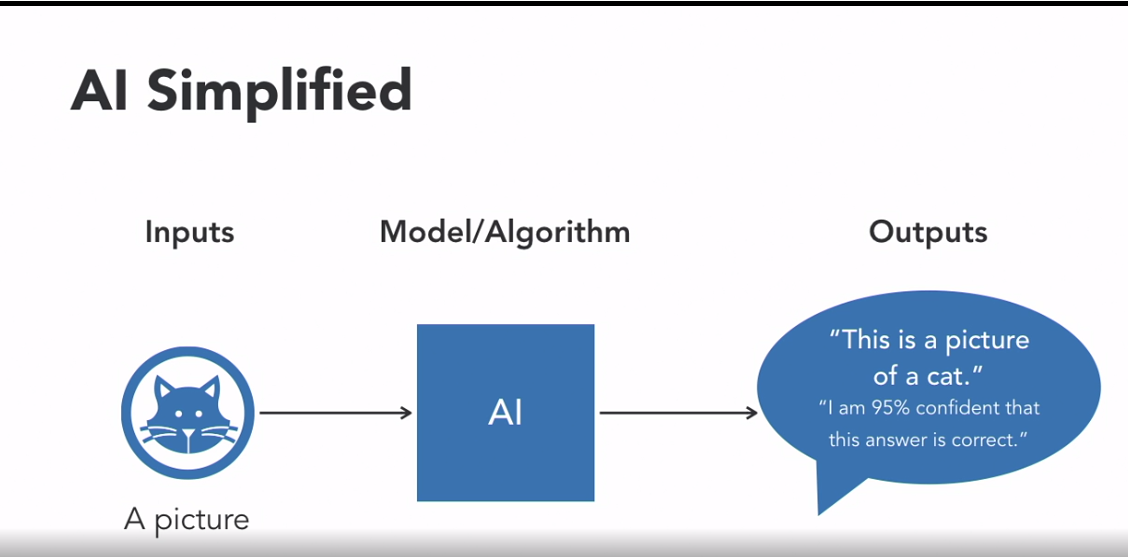


### **Introduction to AI and ML**

Selecting transcript lines in this section will navigate to timestamp in the video

- [Narrator] With the proliferation of artificial intelligence and machine learning algorithms in the world, it is getting more and more important for human beings to be comfortable with using these sometimes mysterious systems. Before we get into the details of XAI, I will take a few minutes to review the basics of AI, machine learning, deep learning, and go over some other terms used throughout this course. Artificial intelligence, or sometimes called machine intelligence, refers to a very broad set of technologies that is developed to make computers or machines intelligent. Now, the term intelligence itself is subjective and can be defined differently, depending on who you talk to. There is certainly a lot of ambiguity about what AI actually is. In fact, some people like to say that a system is only called AI until it becomes used in everyday products. Then, it's just technology. We don't need to worry about these distinctions in this course. It is sufficient to just know that the AI we will refer to here is generally what is discussed in the media and in business situations. Now, let's talk a bit more about what an AI system is. If we massively simplify things, these systems have three main parts. There is the input, there is the model or algorithm, and there is the output. Inputs are the data that you want the AI to analyze. These could be some photos, data from a factory, or sensor information from your self-driving car. The outputs are the decisions the system makes. Continuing with the previous example, it would be the fact that the picture is of a cat, the settings you need to make the factory run smoother, or the fact that your car doesn't run into something on your way home from work in auto-pilot mode. Then, there's the meat of the system: the model, or the algorithm. This is the part that does the actual analysis of the input data to get the output. There are many, many different types, far too many to get into here. Within the large category of AI, there is a subset of methods that uses statistical techniques to learn how to classify or predict outcomes from an existing set of data. This is what's called machine learning. As the name suggests, the system is taught or trained to be able to make these classifications or predictions. Briefly, in machine learning, you would need to prepare a large amount of data, called the training set, to teach the system. We'll talk a little more about the details of this in a subsequent slide. If we go one level deeper within machine learning, we have deep learning, which is based on architecture that is somewhat similar to how our brains work. In fact, deep learning is really just a new name for neural networks, a system that has existed for years. We'll talk a bit about deep learning, as this is often the system that researchers are trying to make more transparent by making it explainable. Here, we have an example deep learning system. You can see that we have an input layer, hidden layer, and output layer. Often, there are many hidden layers, depending on the specifics of an application. In general, the input layer will break down the data it receives into basic elements and then hands them off to the first hidden layer. In the example of a photograph, these basic elements are often individual or small groups of pixels. Then, each set of hidden layers will group these into larger components that it finds. For example, edges of objects, groups or combinations of edges, larger objects, and so on. The final layer, again, is the output layer, which is the decision or recommendation that you want the system to make. In the cat example, it would be the decision of whether the photo is of a cat or not. A big problem with AI systems today is that the process is not transparent. If you asked your friend to perform the same prediction task, she may pick a photograph from a pile and tell you that this is one of a cat and that she is quite certain about that decision. If pressed, she could also explain to you that she came to this conclusion because she identified square ears, long whiskers, and paws that she associates with cats. Unlike humans however, current AI systems cannot explain to you how it came to its decision. In the trivial example of cats, you may not care so much about why some photos were identified as cats and why some were not. However, in more critical cases, such as the recommendation to remove a different organ than originally intended during a surgery, or buying a large position in a stock of a company that you thought did not have a bright future, you may want to know why that recommendation was made. This is where XAI, or explainable AI, would be extremely valuable.

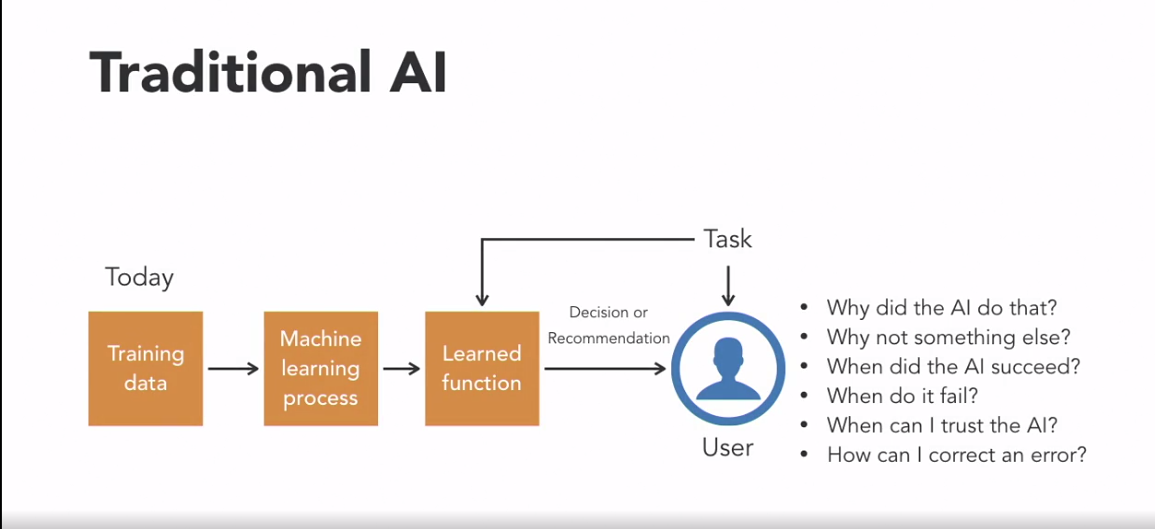


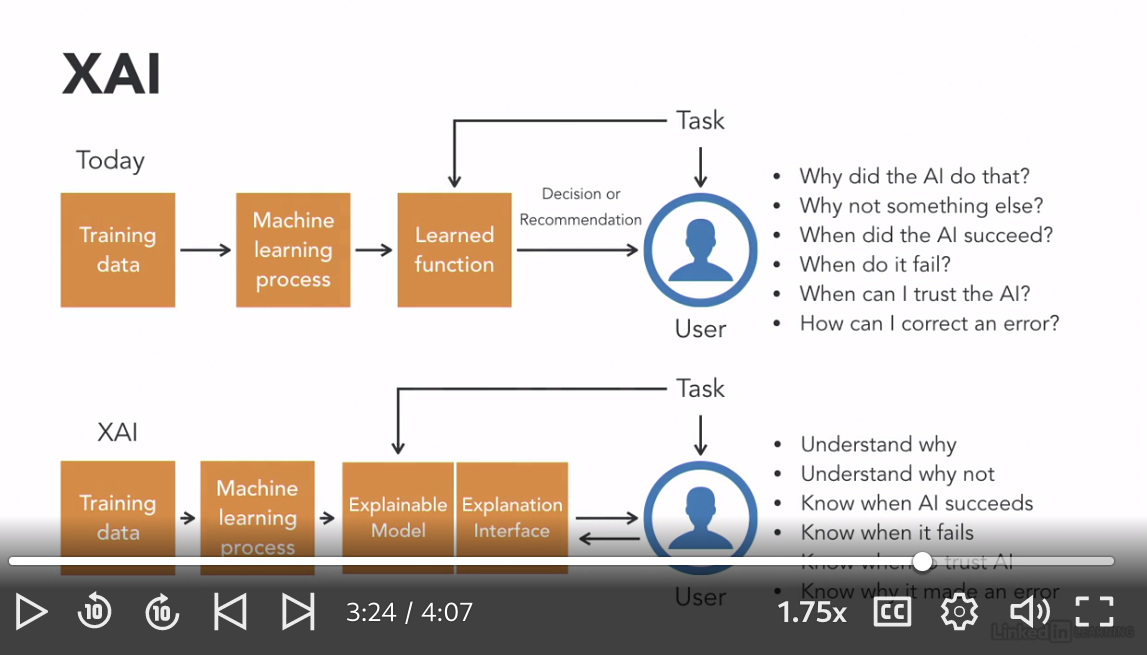


### **What is XAI?**

Selecting transcript lines in this section will navigate to timestamp in the video

- [Instructor] Let's start by talking about why we need XAI. Here's a simple example. I'll start by asking you a tough question. What do you think this is an image of? And if you think you got the last one, what about this one? Okay, those first two might have been pretty easy, but what about this one? And finally, this one? The answers might surprise you. Since this is a course about AI, it would be interesting to ask an AI system what it thinks these images are. In fact, that is what some AI scientists did in some research conducted in 2015. They fed these images into the state-of-the-art deep neural net, a type of AI. The first two answers here were not surprising. The system decided they were images of a guitar and a penguin, much like what most of you probably thought. The numbers shown are not the probabilities of the answer being correct. They are the degree of confidence that the system thinks that it is correct. let's see what the system thought the other two images were of. As you can see, the system also thought the other images were of a guitar and a penguin as well, at extremely high confidence levels. How could this be? In fact, this was a rigged example, in that the data scientists created these two images themselves to trick the system into thinking they were of a guitar and a penguin. We won't get into the details here, but basically, the scientists preserved the elements that the system used to make those classifications and modified the rest of the image. Because these elements are completely different than the ones that you and I use to determine guitar and penguin, we have no idea what these images are supposed to represent. These are trivial cases that wouldn't usually have much real-world impact, but it would be more meaningful to say the least if this was a case where an AI-based system recommended that you should buy a stock that experts said was worthless, or that the black dot on your MRI was cancer. You, or at least your stockbroker or doctor, would probably want to know why the system made those recommendations. How would an explainable AI-based system actually work? The first thing to know is that the research is still in its infancy. There are many research efforts going on around the world. Some of them will probably succeed, but it's likely that far more of them will fail. Let me explain the high-level concept of how this might work. In the normal process of a typical AI system, there are a few steps. First, training data is identified. Then, it undergoes the machine learning process. Then, a learned function or algorithm is produced. Through this algorithm, the system can make decisions or recommendations to a user for a specific situation. However, the user will not know things like, why did the AI do that? Why not something else? When did the AI succeed? When does it fail? When can I trust the AI? How can I correct an error? As can be seen in this diagram, an XAI system could replace the traditional learned function with an explainable model. This model would be built in a way where the decision-making process could be understood by human beings. This process could be presented to the user through something like an explanation interface. As a result, the user could understand why, understand why not, know when the AI succeeds, know when it fails, know when to trust the AI, and know why it made an error. XAI makes the black box, opaque nature of AI a little more transparent. In the following video, we'll cover three of the more common XAI techniques being worked on today.

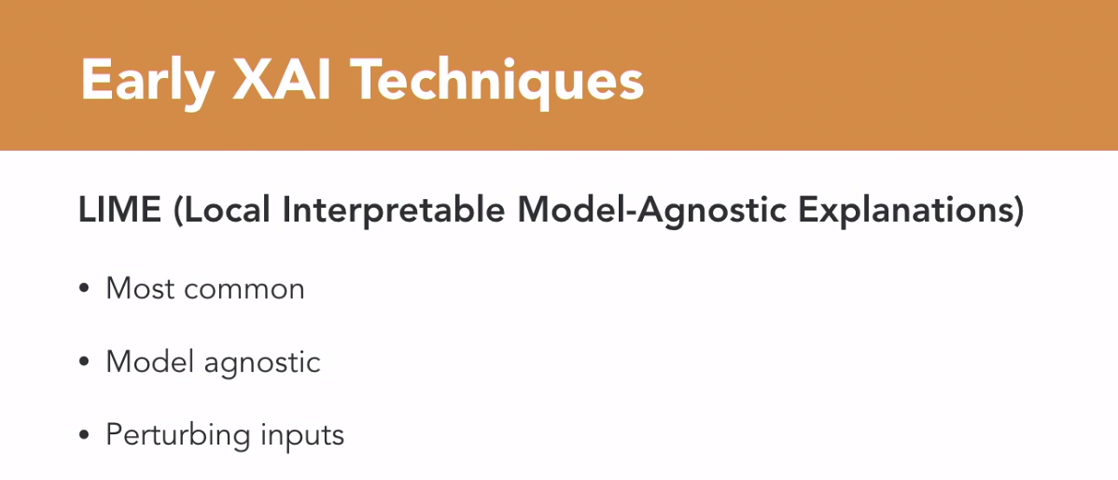


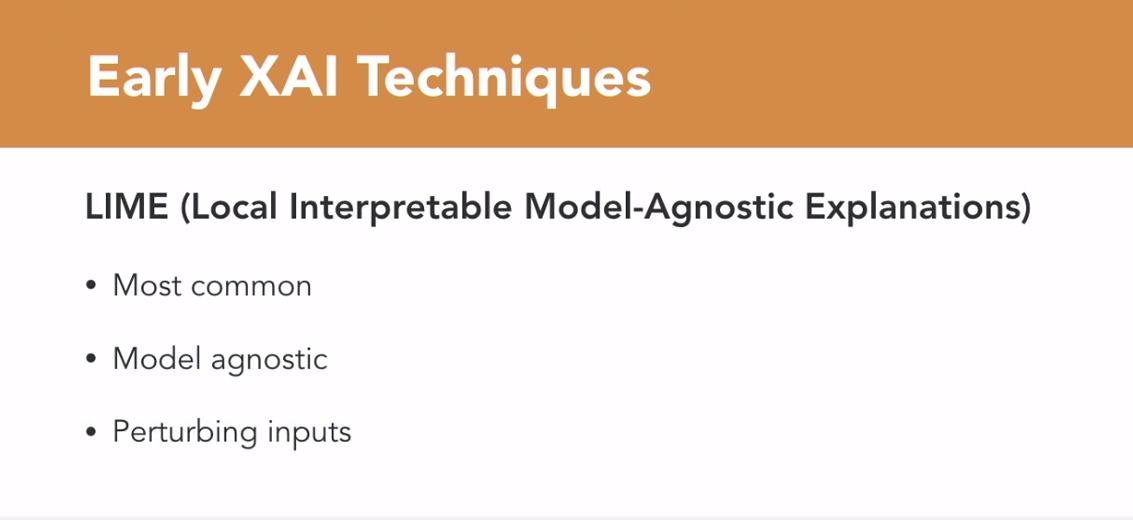


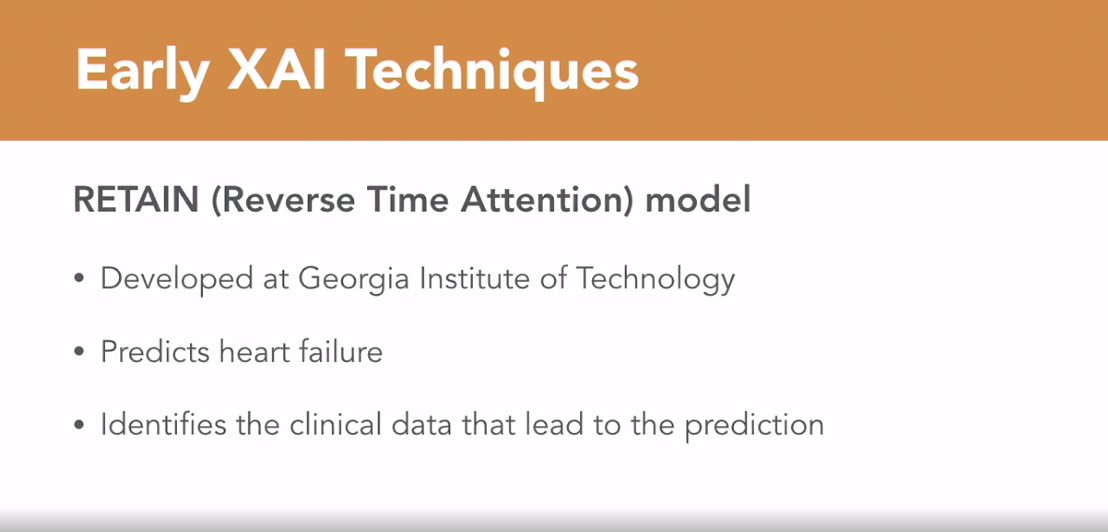
### **XAI techniques**

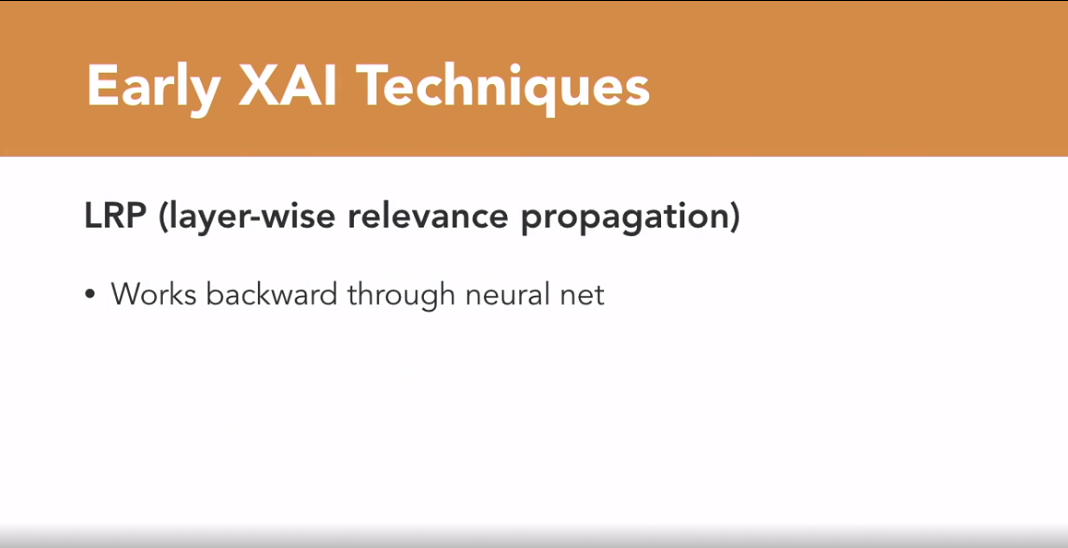
Selecting transcript lines in this section will navigate to timestamp in the video

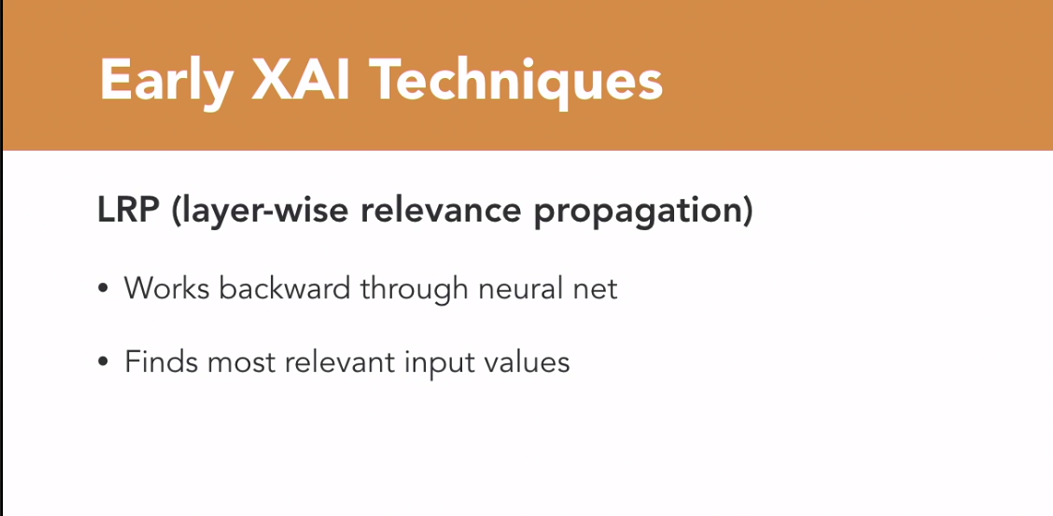
- [Instructor] Though the field is early and rapidly changing, three of the more advanced XAI techniques are **LIME, RETAIN, and LRP. Currently, the most common XAI technique is LIME.** It stands for **Local Interpretable Model-Agnostic Explanations**. LIME is actually a post hoc model, which means that it is a technique that looks for an explanation after the decision has been made. One benefit of LIME is that, as its name states, it is model agnostic, meaning it doesn't matter what type of model it is applied to. The LIME approach involves perturbing or slightly changing the inputs of the model to observe how the outputs change. This allows us to understand which inputs affect the outputs the most, giving us insights into how the model made its decisions. RETAIN, which stands for Reversed Time Attention model, was developed at the Georgia Institute of Technology to study models that predicted heart failure. In this method, the model is designed such that, using the data from patient clinical visits, it can predict the occurrence of heart failure at a rate comparable to other models, but also identify which piece of clinical data contributed to the prediction. LRP, or layer-wise relevance propagation, works backwards through a neural network and figures out which input values were the most relevant in coming up with the output. Though these are three common XAI techniques being worked on today, there will no doubt be many others to come as the field develops. In the next two videos, we'll talk about why you might need XAI from the business and legal perspectives.

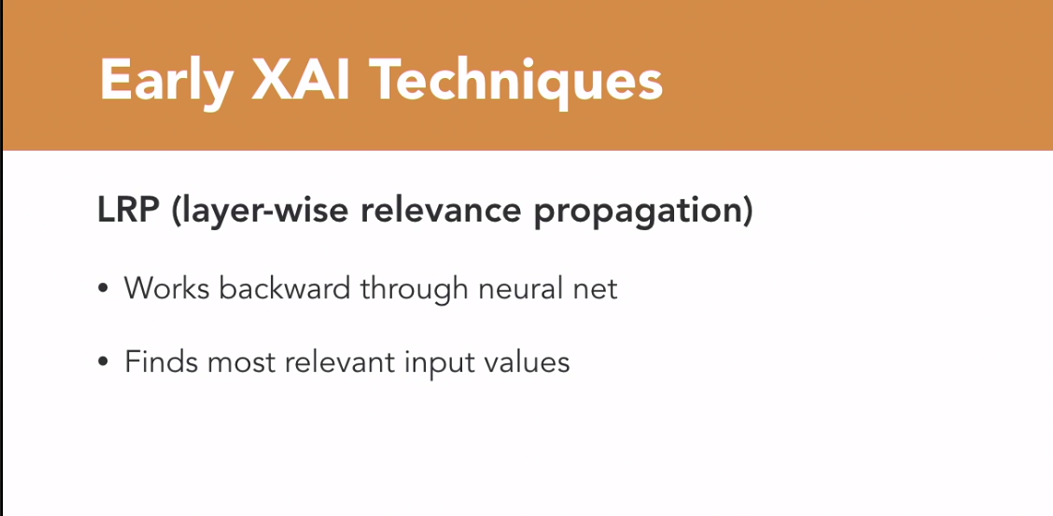








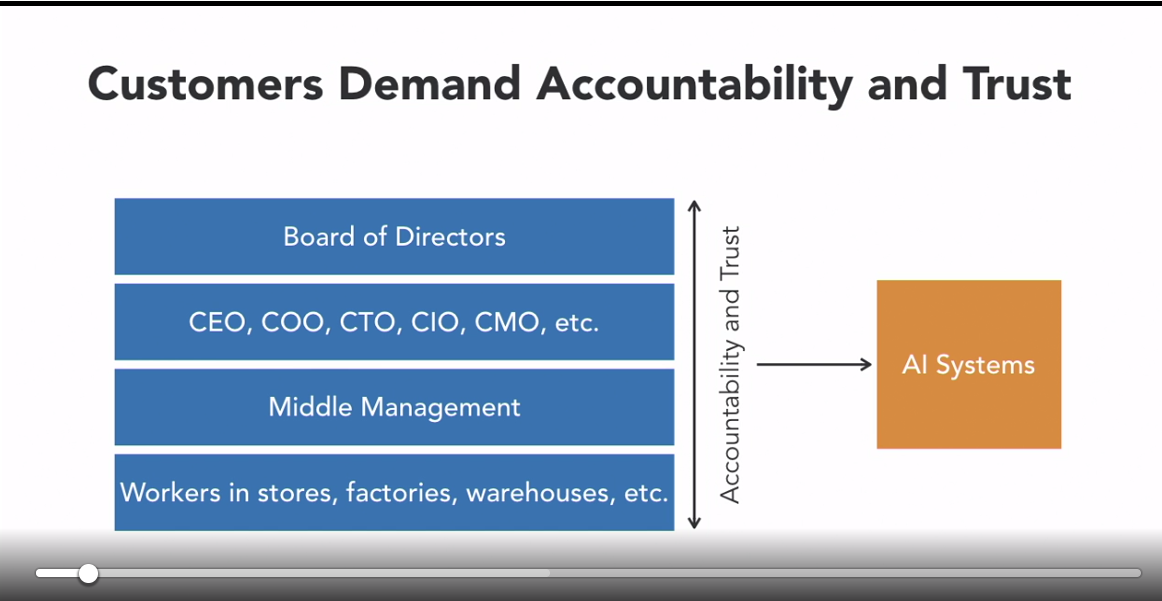


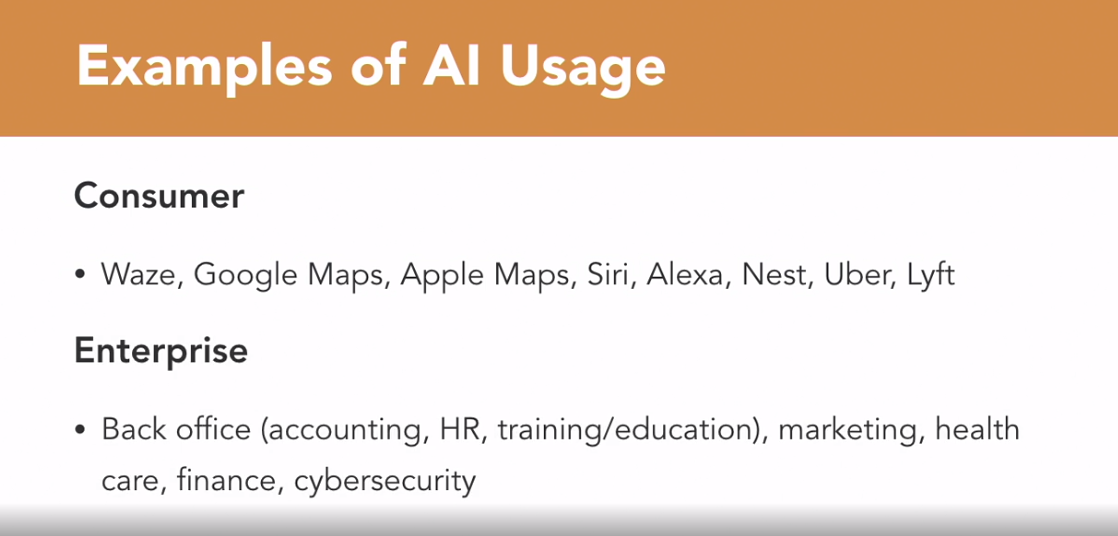


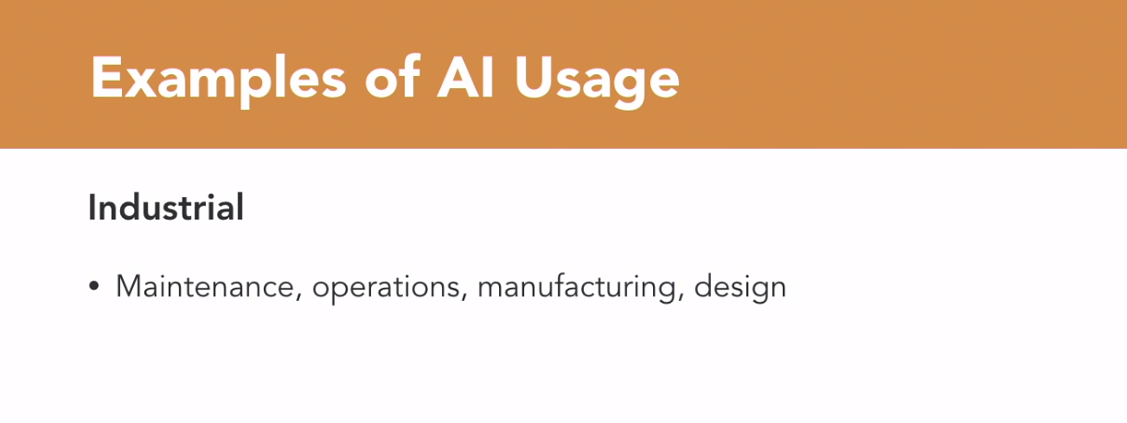
### **The need for XAI: Business**

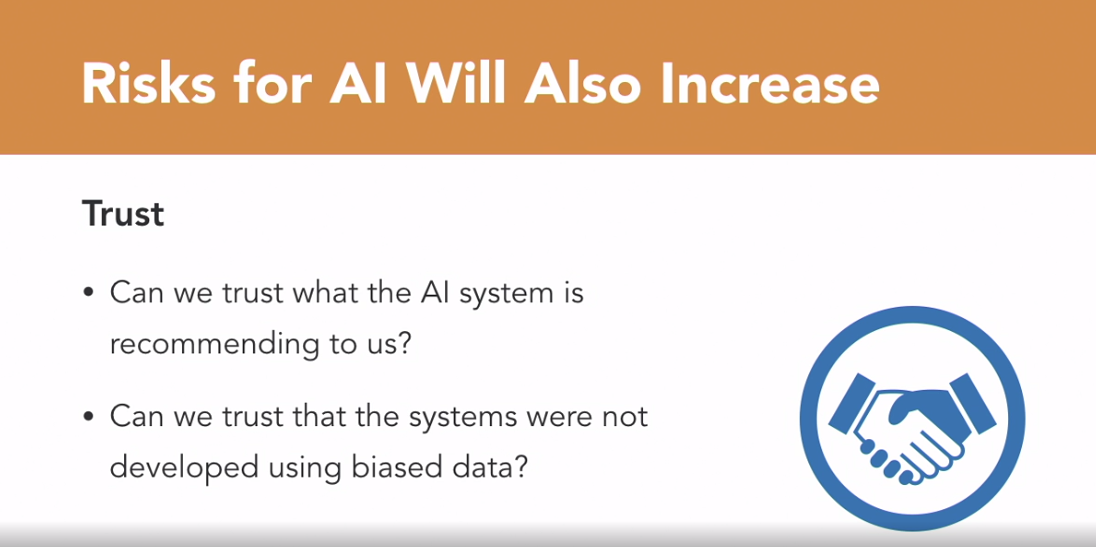
Selecting transcript lines in this section will navigate to timestamp in the video

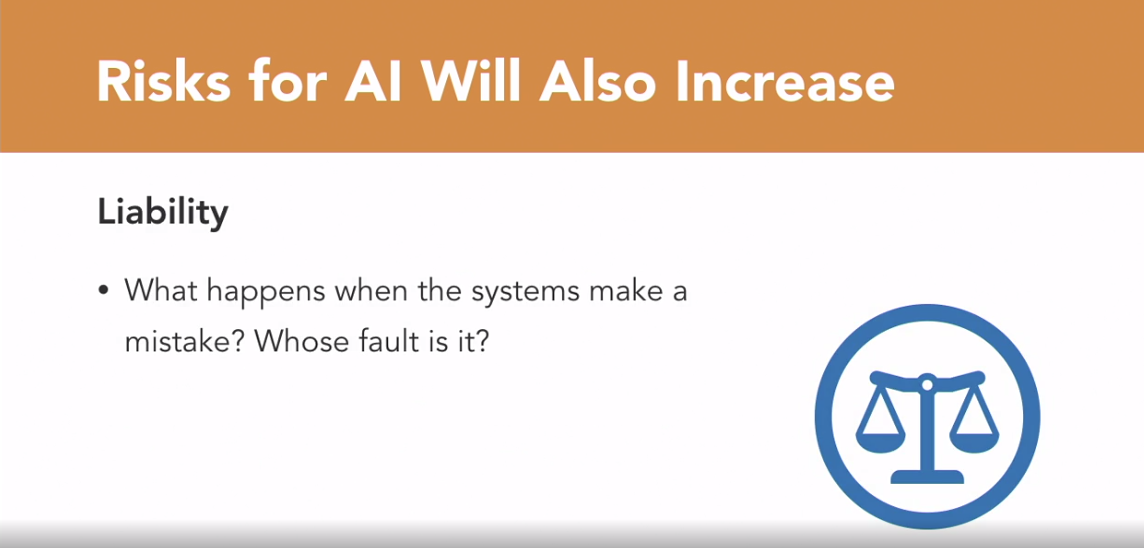
- [Instructor] In this chapter, we will talk about why the explainability of AI systems will be important for businesses. Two of the core tenants of any successful business are accountability and trust. Customers and partners expect this from frontline workers in the store, factory, or warehouse through middle management all the way to the CEO and the board of directors. As AI systems start making more and more decisions that humans used to make, these tenants will be demanded of these systems as well. AI is now all around us. For consumer applications, it is embedded in things like Waze, Google Maps, Apple Maps, Siri, Alexa, Nest, Uber, Lyft, and so on. In the enterprise, it is already used in back office systems such as accounting, HR, training and education systems, as well as for marketing, health care, finance, cybersecurity, and a whole host of other applications. For industrial applications, maintenance and operations systems, manufacturing and factory control as well as design systems all incorporate AI. The penetration of AI into the many facets of our lives will only continue to accelerate over time. As AI becomes more ubiquitous in our lives, there will be more and more areas of risk for AI that will need to be addressed. Examples include trust. Can we trust what the AI system is recommending to us? Can we trust that the systems were not developed using biased data? Liability. What happens when the systems make a mistake? Whose fault is it? Security. How can we know if the systems have not been maliciously manipulated? Control. Who, machine or human, has control of a process? If machine, then how can a human take it back? Because the consequences for inaccurate recommendations are minor for consumer applications, AI adoption has been more rapid. If your smartphone recommends a particular route to your house over another or Amazon recommends a certain product over a different one, the implications whether positive or negative are relatively trivial. For decisions that have large financial implications, impact a very large number of people deeply, or have life or death consequences, adoption of AI has been slower. XAI may alleviate concerns over transparency, bias, and reliability, and speed up the adoption of AI in these areas. For example, what if your AI based medical system diagnoses some black spots on your skin to be cancer. Would you take that at face value and elect to have surgery or chemotherapy right away? Probably not. You likely would not 100% trust that decision. Currently since you can't ask your AI doctor why it made that diagnoses, you would ask your human doctor. In addition to increasing the accountability and trust of AI systems in general, XAI will also help scientists and engineers improve the performance of AI systems within business. An explainable system will be easier to debug for unintended biases that may result from skewed training data. It will be easier to detect malicious hacks of explainable systems. Explainable systems will be easier for humans to work with thus leading to man machine collaborations that will perform better than just computers or humans alone. There will be a more detailed example of this in chapter two, the centaur chest example. Here is another example where a business might benefit from XAI. A system may tell the agent how it made the decision to approve or deny loans to clients. The agent can review the logs and decision reports to make sure there are no inappropriate biases caused by skewed input data. XAI will help businesses with more transparency, accountability, accuracy, reliability, and so on which will create differentiation, better decisions, and more trust and ultimately help companies be more successful in their markets.

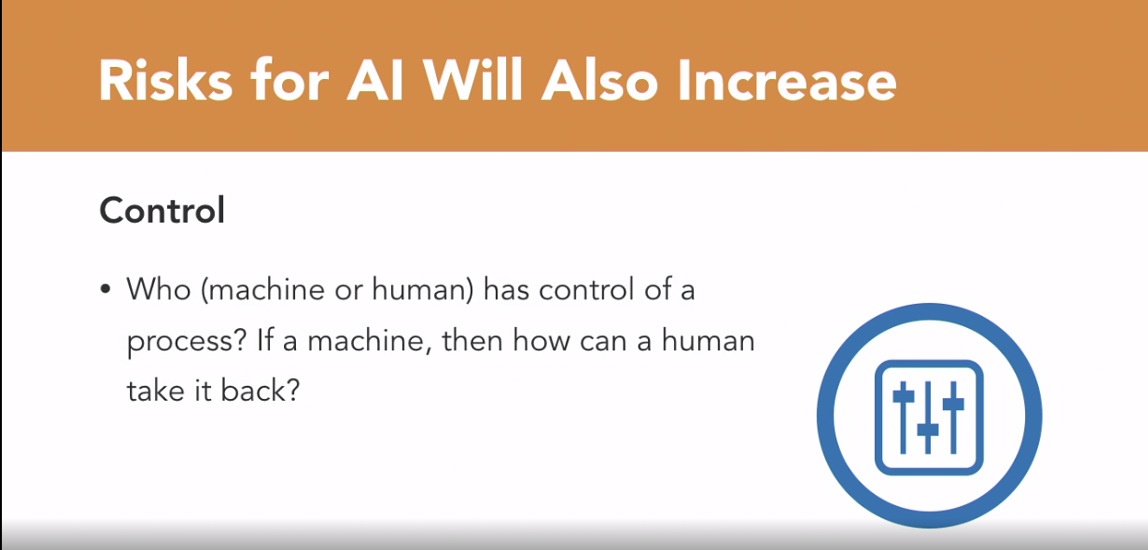


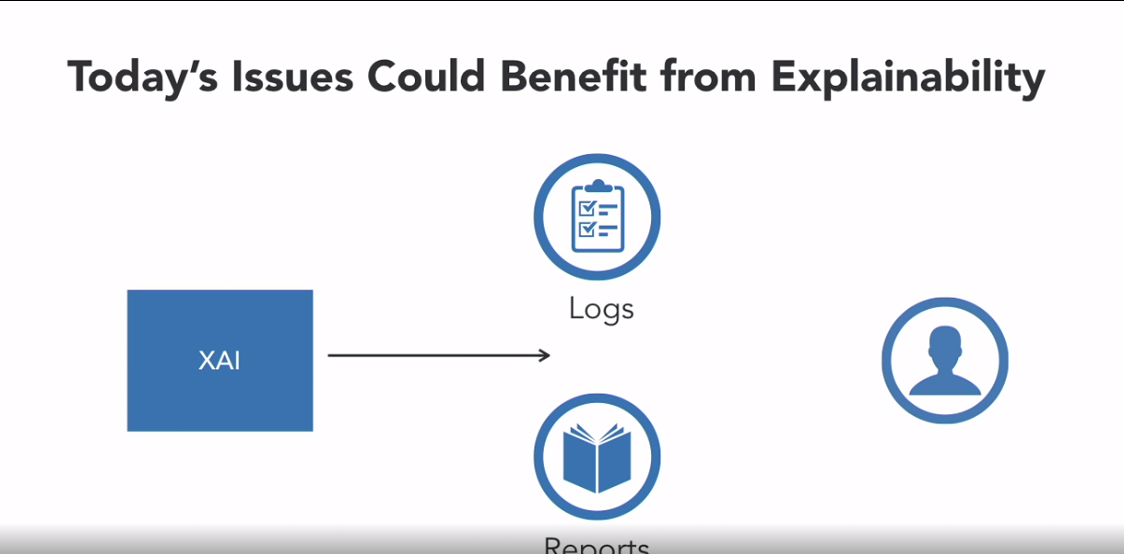












### **The need for XAI: Legal**

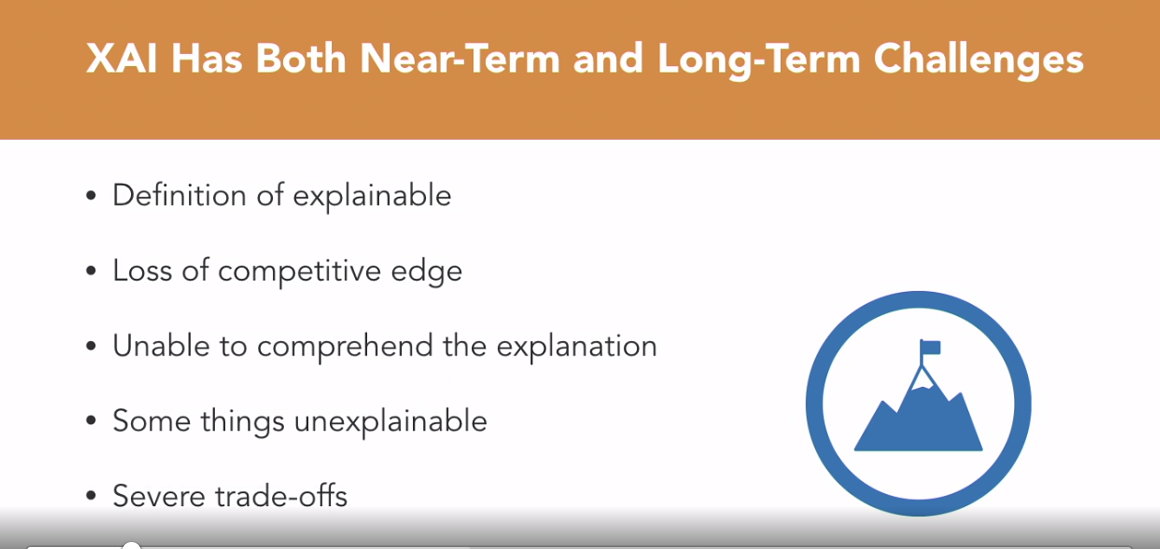
Selecting transcript lines in this section will navigate to timestamp in the video

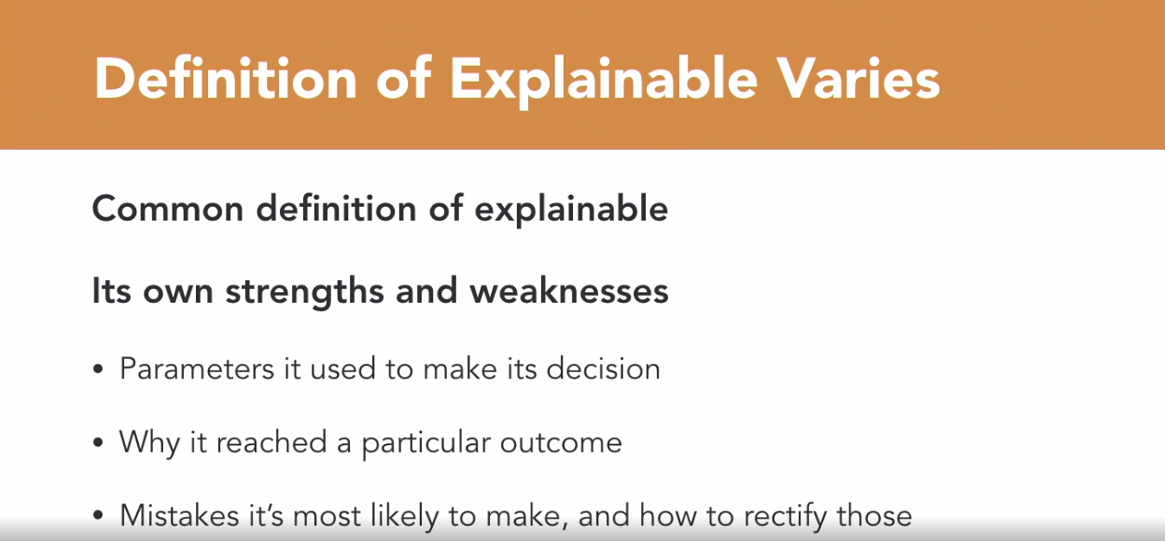
- [Instructor] In the previous video, we learned about why XAI might be important for the success of your business. In this section, we will look at the legal reasons why you might need XAI in your business. Currently, the main driver for this is GDPR. The GDPR, or the General Data Protection Regulation, is a European regulation focused on data privacy and protection for individuals that became effective in May of 2018. Fines for noncompliance can be as high as 20 million Euro or or 4% of annual revenue, whichever is higher. It can be quite costly for businesses that are caught not fulfilling the obligations of GDPR. The section of most interest to us related to XAI is Article 22 of the GDPR, often referred to as the Right to Explanation. The GDPR Section 22 states, "The data subject shall have "the right not to be subject to a decision based solely "on automated processing, including profiling, "which produces legal effects concerning him or her "or similarly significantly affects him or her. "the data controller shall implement suitable measures "to safeguard the data subject's rights and freedoms "and legitimate interests, at least the right "to obtain human intervention on the part of the controller, "to express his or her point of view "and to contest the decision." while the detailed requirements of this section in the GDPR is still being hashed out by the politicians and lawyers, there will no doubt be many implications. It would require the disclosure about data, such as what was collected, when it was collected, how it was used, who it was shared with, et cetera. Also, there is a growing consensus among data scientists that it will generally lead to requiring more transparency and accountability in AI systems. Three such points are often cited. The AI system must allow for human intervention. There must be an ability to obtain further information from an AI system. The subject has a right to contest a decision made by an AI system. This may affect you in two ways. If your business is serving users in Europe, you will be directly required to comply with this law. However, even if your business is not aimed at European customers, it is possible that similar laws will be enacted in your region as well. The GDPR is not the only law with implications for XAI. Though not as far-reaching, the Equal Credit Opportunity Act passed in 1974 as wording that requires credit-rating agencies to notify consumers of the main factors related to determining their credit scores. As we will see later, companies like FICO, one of the big US credit-rating agencies, are already taking steps to address this for AI. The French Digital Republic Act of 2016 requires that all decisions made by public sector bodies about individuals need to come with explanations. If, or when, French public sector bodies begin to incorporate AI into their decision making, they will likely need to use systems that are explainable. In addition to general laws like GDPR, the Equal Credit Opportunity Act, and the Digital Republic Act, industries such as healthcare, finance, legal, and automotive may have regulations that require a large amount of transparency. The adoption of AI in some of these industries is already starting to slow because of its black box nature. For example, before banking and financial institutions are allowed to use more sophisticated AI systems that can move millions or even billions of dollars without human oversight, they will likely need to make these systems more transparent or explainable. XAI will make these systems more transparent, accountable, and accurate, which will allow regulators and users alike to trust them more. In addition to the business benefits for XAI, many industries will be required to use systems that are explainable for legal reasons. The most impactful law currently is the GDPR in Europe, which is already starting to change the way companies work with data. Countries will also most certainly continue to pass new laws that will require companies to be more open and transparent with algorithms. Finally, industries that are highly regulated will be one of the first to need systems like XAI. However, before we continue to talk about the benefits of XAI, let's briefly touch on some of the challenges and limitations of XAI in the next video.

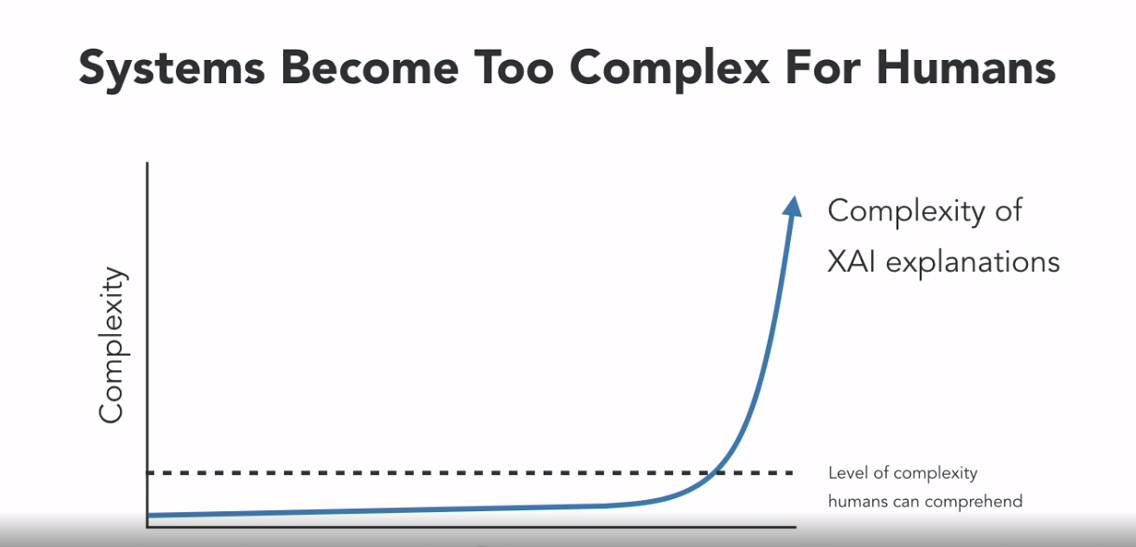
### **Limitations of XAI**

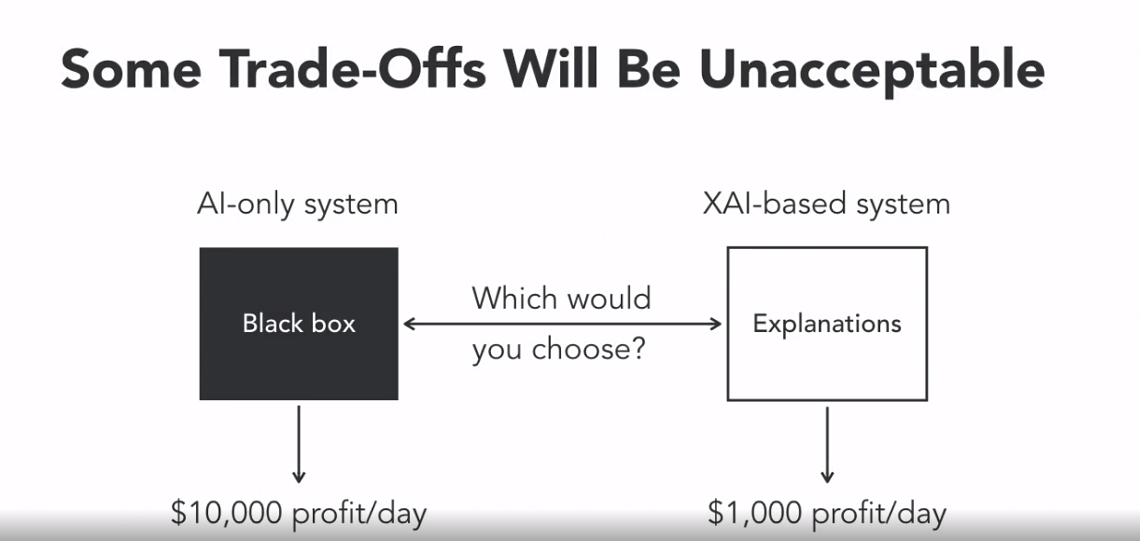
Selecting transcript lines in this section will navigate to timestamp in the video

- [Instructor] Although we talked about the many benefits of XAI, it is important to note that there are some significant challenges for the technology both in the short term as well as into the future. In this video, we will cover the difficulty in defining explainable, the potential loss of competitive edge for companies that incorporate XAI, the difficulty in understanding the explanations as they get more complex, the fact that some things are inherently difficult or impossible to explain, and the trade offs that you might have to make to achieve explainability. First, the very definition of explainable is still in question. Explainable can mean everything from understanding the detailed actions of the AI algorithm to higher level concepts of reasoning and logic. The current consensus among data scientists about the meaning of explainable seems to be that the system will tell you about its own strengths and weaknesses, the parameters it used to make its decision, why it reached a particular outcome as opposed to a different one, and the mistakes it's most likely to make and how to rectify those potential mistakes. However, to what level of detail does it need to go to to be considered explainable? Some people have stated that justifiable, responsible, or transparent are better terms than explainable as it implies deeper levels of logic and reasoning. As you can see, there are many different questions about the meaning of explainable. Another potential issue for using explainable AI is that it may compromise a business's competitive edge or differentiation. For example making complex AI systems easy to understand by non-technicians or the general public may mean that trade secrets and competitive differentiation could be compromised. Understanding how a black box works may allow it to be copied or replaced by other solutions that perform similar functions. If the advancement of AI systems continues, it is conceivable that these systems will at some point become so complex that even the simplest explanations may be too complicated for the smartest humans to understand. Take the analogy of a nuclear scientist trying to explain the details of a complex differential equation to a kindergartner. Even thought the scientist and the youngster are speaking the same language and understand all of the words being spoken, the concepts would be just too complex for the five year old to understand. While this may seem to be science fiction or thousands of years in the future, with geometric if not exponential advancement of technology, this could happen much faster than most people expect. Another problem with the concept of explainability is that some things are just not easily explained. If you think about it, humans are actually not that good at explaining some of their decisions. We often make decisions based on emotion or non-rational factors and then reverse engineer those decisions to come up with explanations. Some experts have said that in many cases, humans make decisions based on emotion, intuition, or other irrational reasons and then they construct logical explanations after the fact that justify those decisions. You may be able to logically and rationally explain why you like pizza, why you picked the color of the car you bought, how you ended up with your spouse, but chances are that a large part of the decision is based on emotion or other factors that may be hard or impossible to explain. If we are not able to explain our own decisions, some may argue that we should not expect computers would be able to explain their decisions either. There will likely be unacceptable or undesirable trade offs in achieving sufficient explainability. These trade offs could result from the fact that some approaches for creating XAI systems will involve rethinking how AI systems are developed. While some of these approaches may increase the output quality of the systems, we can expect that other approaches may lead to some degradation in the results due to changes in the algorithms that will be necessary for humans to comprehend how they work. XAI may not be the solution for every system. Some of the trade offs of having XAI may be minor but others might be more major. For example on one hand you might have a stock trading system that could make you $1,000 a day with XAI but in doing so, it is inherently encumbered by the requirement to explain every decision. On the other hand, you might have another system that was making you $10,000 a day on average but was a complete black box, a much more efficient system that doesn't need an explanation for every step. It would be safe to say that many, if not most would choose to make $10,000 a day and forgo the explanation regardless of the added transparency. As you can see, in addition to the technical challenges of actually building a meaningful XAI system, there are many other challenges and limitations of XAI that need to be resolved. It would be good to keep these limitations in mind as the technology continues to develop.





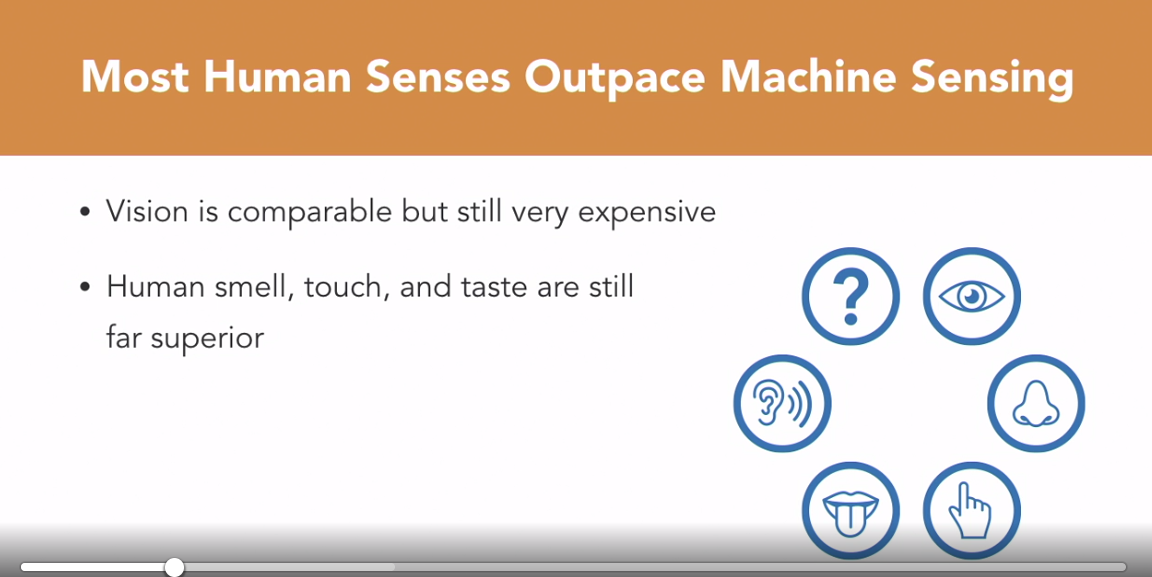


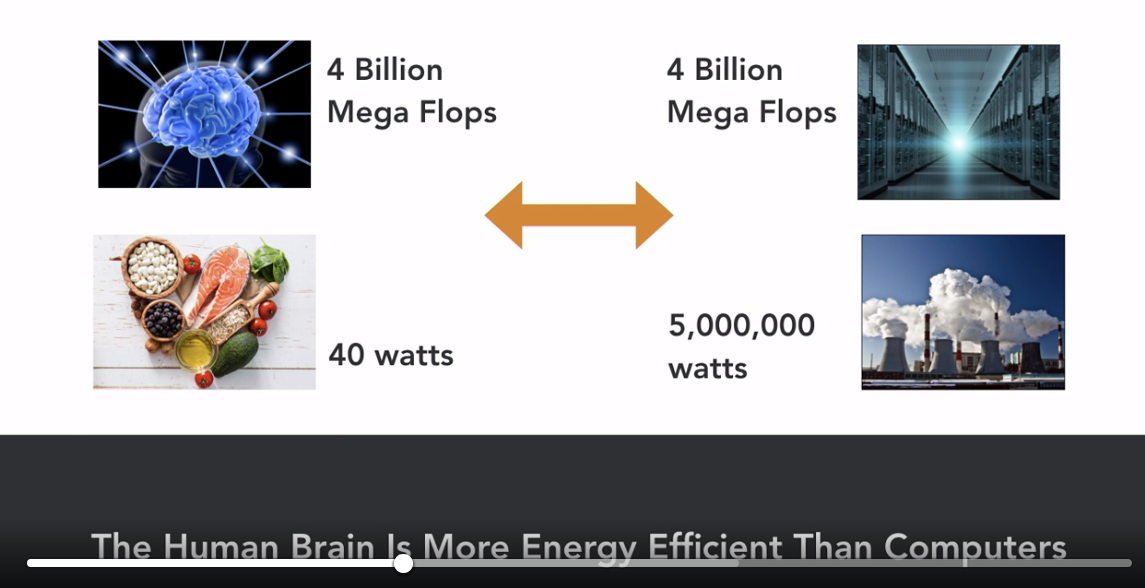


### **Humans are better at some things**

Selecting transcript lines in this section will navigate to timestamp in the video

- [Instructor] I'd like to illustrate why it makes sense for humans and computers to collaborate with each other. To start we'll discuss some things that humans are better at and then some times that computers are better at. Finally, we'll talk about how XAI could make human machine collaboration much better than it currently is. So let's get started on some examples of things that humans are better at than machines. First, let's look at some of the sensing abilities of humans. Though expensive machines now have the ability to see the world at the same resolution and same throughput speed as the human vision system they are not yet widespread. The human nose however is still far superior to any mechanical smelling mechanisms out there. Finally, the ability of the skin to sense things like touch, fine temperature differences, and pain, and the tongue to taste will still be unmatched by machines for a long time. Humans are still far more creative than computers. Although recently there have been a few cases of computers making a painting or composing a song, by and large humans are much more proficient in the creative fields, such as design, art, architecture, music, writing novels, movie-making, playwriting, creating new recipes and so on. Next, let's look at the energy required for processing information. Currently it is estimated that the human brain is about 1,000 times more efficient than a computer. To illustrate let's look at some numbers. It is estimated that the human brain uses about 40 watts of energy to complete about four billion mega flops. Note that one mega flop is a million operations per second. The best super computers today need to consume five million watts to complete the same four billion mega flops. While it may surprise you that memory is on this list, take DNA for example. Every single one of our six trillion cells holds all of the information necessary to create your entire body. This is more efficient and compact than any hard drive. In addition while digital media degrades over time, there has been DNA from Neanderthals dating back 30,000 years ago that are intact. DNA from other species from over one million years ago have also been found preserved. In addition, humans are much superior to machines in fine motor skills. Take your hand for instance. It is composed of 27 bones, all finely connected and activated by numerous muscles and tendons. These all combine to be able to manipulate everything from threading a needle, juggling, to playing thumb wars with each other. It will take a long time before mechanical systems will be able to work so well together. In addition to the physical aspects, humans are still better than computers at processing certain types of information. Unstructured problem solving or problems where the rules are not yet clearly defined is one of these areas. For example, in a factory that has hundreds of machines and multiple ways each machine can be configured you will have a massive number of ways these can all be combined to produce products. Using experience and intuition an expert engineer could find the best process to develop a new product rather quickly. It would take a machine many hours, if it could do it at all, to do the same thing. Humans are much better at selecting and processing relevant information from a field of massive new information. For example, if a reporter goes to report on the elections of a country that has just developed democracy she will be able to better identify and report on topics of most relevance. Humans are much better at other things, like tasks that consist of non-routine physical work, such as cooking different meals every day for your family, or collaborating with many other humans, which happens in every medium to large company. Let's look at a graph with the horizontal axis representing the complexity of a problem and the vertical axis representing the amount of effort necessary to solve that problem. Initially as a problem gets complex it will take a lot of effort for a human to solve that problem, however as the problem becomes much more complex humans can start taking advantage of intuition and experience or heuristics to solve them. For example, summing 10 sets of three to four digit numbers is doable, but not easy for a human being. However, the complex task of planning a two week European vacation for a family of four is also hard, but certainly manageable. In other words, the curve is pretty steep initially, but levels off as you move right. In summary, thankfully there are still a huge number of things that humans are better at than machines. In the next video we'll look at some things that machines do better than humans.

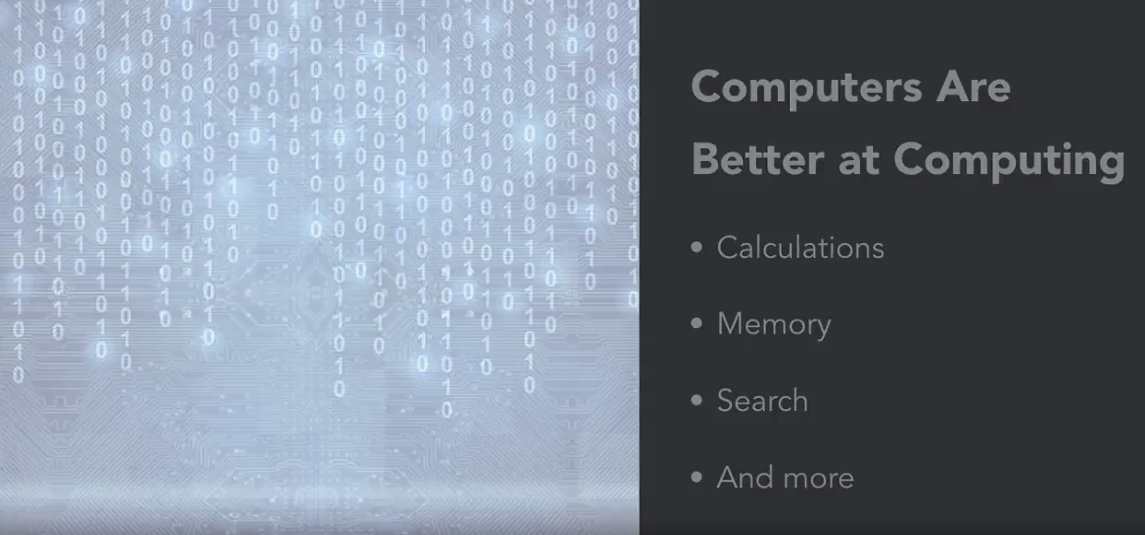


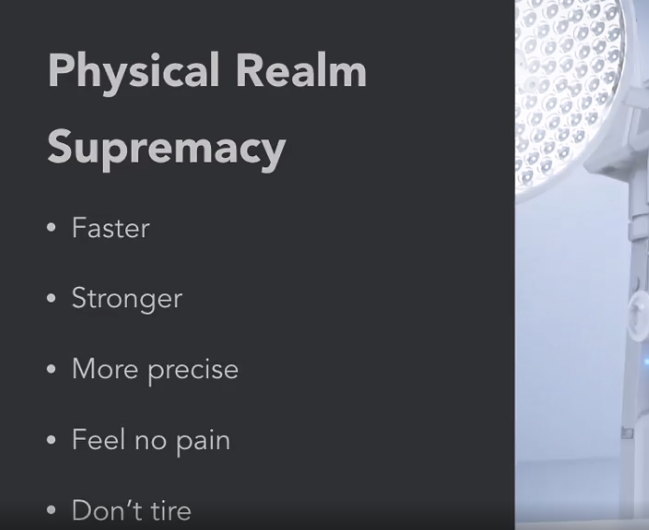


### **Computers are better at some things**

Selecting transcript lines in this section will navigate to timestamp in the video

- [Instructor] In this video, we'll now look at some areas where machines are better than humans. They are pretty good at games. Computers beat the chess grandmaster Garry Kasparov in 1997, and in the much more complex game of Go, one of the most dominant human players, Lee Sedol, in 2016. In 2011, IBM's Watson beat Ken Jennings, who had the longest unbeaten run at 74 winning appearances, and Brad Rutter who had earned the biggest pot of $3.25 million. In the processing realm, the list is quite long, and getting longer all the time. Machines are better than humans at things like, numeric calculations, remembering massive amounts of facts without error, searching through extremely large amounts of information, and so on. Specifically, with the recent progress made in deep learning as discussed in the intro video, AI has surpassed humans in many image processing tasks, such as identifying a large number of objects, like faces in a crowd, or detecting product defects on the factory line quickly and in real-time. In the physical realm, machines are far superior to us in many aspects. They are generally faster, stronger, more precise, feel no pain, and never tire. Other advantages for machines are that they are replaceable, replicable, reliable, obedient, and predictable. In fact, they have already started replacing people in entire jobs where tasks are repetitive and predictable especially in dangerous or undesirable jobs. For example, machines are doing some of the following tasks that humans used to do, factory worker, switchboard operator, bowling alley pin setter, bridge toll collection, grocery store cashier, ticket sellers of all types, just to name a few. It is predicted that they may even begin replacing or heavily augmenting tasks of positions such as lawyers, photographers, customer service representatives, caregivers and drivers. Going back to the solution effort versus problem complexity chart from the previous video, you can see that although computers can solve math problems with virtually no effort, complex problems where success can vary, such as planning a vacation, are next to impossible to solve. In summary, in these two past videos we have discussed things that humans are better at, and things that computers are better at. Humans are generally better at making sense of things in very complex undefined environments. Most of the creative fields, such as art, music and design, are still ours. Other than sight and sound, our senses are much more advanced, and we're able to better perform tasks that are non-repetitive and undefined. Machines, on the other hand, are better at defined, bounded problems, like mathematic calculations, remembering facts and figures, and a growing set of image processing tasks. In the next chapter, we'll bring these two things together to talk about how human and machine collaboration can be much more powerful than just one or the other alone.





### **Example: Centaur chess**

Selecting transcript lines in this section will navigate to timestamp in the video

- [Instructor] In the previous video, we made the argument that it is good for humans and machines to collaborate. Now, let's go through a specific example of where this collaboration, when done well, is better than all other combinations. The example we will use in this case will be based on chess. Chess is a game that is played with 32 total pieces, 16 for each player, and is estimated to have at least 10 to the 120th power possible moves, which is far more than the number of atoms estimated to be in the universe. The sheer number of possible moves makes it so complex that even computers today have not been able to calculate all of the different outcomes, which makes it, as yet, an unsolved game. This makes it fascinating for computer scientists to use for AI. Since about the 1950's, programmers have been developing software to play chess. At first, progress was good but relatively slow. But with the development of technology accelerating faster and faster, these systems continued to make solid progress over time. At the same time, the skills of human chess players developed much more slowly, if at all. It was inevitable that computers would be better than humans at chess someday. That day came on May 11th, 1997 when IBM's Deep Blue defeated Gary Kasparov, the first time an AI beat a chess grand master. And since the curve for AI improvement is much steeper than for humans, we are no longer in the race. As a point of comparison, today's mobile phones are many times more powerful than Deep Blue was in 1997. Compressing the time line down a bit, you can see that there is no hope for humans at ever beating the top computers at chess again. This then brings us to the conclusion that as of today, a strong computer is much better than a strong human at chess. However, Kasparov suspected that humans might have been helping Deep Blue in his match. We don't know if he ever validated that suspicious, but this led him to wonder, what would happen if humans and computers worked together? From that idea, in 1998, the year after his big defeat, he held a freestyle chess tournament where teams of computers and humans called Centaurs could play each other. Kasparov wanted to know if a human in the loop could actually help a computer beat a grand master in chess. Time went on and in 2005, inspired by Kasparov's idea, a online freestyle chess tournament was held by Playchess.com. Most online tournaments barred the use of computer aided play. Playchess wanted to see what combinations would be the best and allowed all comers. Competitors included Hydra, a super computer much more powerful than Deep Blue, as well as many types of other chess software, multiple human grand masters, and other human amateurs, and a combination of the above. Not surprisingly, the addition of a computer to a strong human competitor was better than just a strong human alone. Since even a strong human is far inferior to a strong computer, the logical conclusion might be that adding a human to a strong computer would make that combination less effective than a strong computer on its own. For example, an elite tennis player would not be helped by adding an amateur tennis player onto the court for her finals match. However, that did not turn out to be the case. A combination of strong humans and strong computers consistently beat out strong computers on their own. Let's try to take a deeper look at why this might be so. Computers are inherently better at things like logic, computation, memory, and are consistent. Humans on the hand, are better at things like intuition, creativity, design, and finding analogies. Thus one could summarize that with these strengths, the computers are better at humans in choosing answers, humans though, can be thought of as having better skill in choosing the right questions. It turns out that this is a powerful combination when trying to calculate the next move in a game as complex as chess. In fact, not only was the winner of the tournament a human and computer combination, it was a team of two young amateurs working with three weak computers. The three computers ran different types of chess software and when recommendations were different between the machines, the humans interacted with the computers to analyze the moves further. Being able to combine the strengths of humans and the strengths of computers, which we will call superior process, was what allowed them to beat Hydra, one of the most powerful and sophisticated chess programs of that time. In fact, this combination also beat out other grand masters that were also using computers, but using inferior processes. To summarize, it can be concluded that for certain applications such as chess, a superior process trumped superior humans and superior computers and their combinations. There have been other examples such as in fashion design, design of a quad copter body, and part where these superior processes have yielded results not achievable by humans or computers alone. In these examples, the superior processes were developed through experience, practice, and trial and error. However, many AI scientists now believe that better insights by humans into the workings of computers can lead to even more powerful combinations of humans and computers. Explainable AI systems that can show the human user how it is making its decisions could help build teams of man and machine that could solve some of the biggest and most pressing problems that the world faces today.