Generative AI,

How finding and sharing information online has evolved
Selecting transcript lines in this section will navigate to timestamp in the video
For years, we've navigated the vast seas of the internet using search engines as our compass, typing inquiries and scrolling through results. But reasoning engines like ChatGPT and Bing chat have forever changed this journey. No longer are we merely typing keywords. We're having conversations, but this new frontier requires careful navigation. Thoughtful questioning and clear context are essential to achieve meaningful and accurate results. I'm Ashley Kennedy, Managing Staff Instructor at LinkedIn Learning. In this course, I'll be speaking with two top experts in this emerging field, Noelle Russell, who leads the Global AI Solutions at Accenture. And Brandie Nonnecke, Founding Director of CITRIS Policy Lab and Research professor at UC Berkeley. Together, we'll explore the exciting evolution of thoughtful online search.

How a search engine works

How a search engine works
Selecting transcript lines in this section will navigate to timestamp in the video
You've probably used a traditional search engine like Google or Bing countless times for everything from looking up basic information to performing deeper research, to getting news headlines, to shopping, travel, entertainment and more. There's so much you can find out using simple keywords and queries, but if you ever wondered what was happening under the hood to provide you with your answers. A traditional search engine works by performing three main functions, crawling, indexing, and ranking. Let's talk about this. First, crawling. search engines use computer programs, called web crawlers or spiders, to discover new and updated webpages. These crawlers systematically browse the internet following links from one webpage to another. This process of crawling is continuous, and its goal is to keep the search engine's index up to date. Next, indexing. Indexing takes all the content found during the crawling process, and stores, and organizes it in a massive database called the search index. This is where a search engine keeps all webpage information, and also analyzes various factors like keywords, metatags, headings, and links to understand the relevance and context of all of this data. Then, there's ranking. Ranking is the process of providing the best, most relevant results in response to a query. That means each time you enter a query, the search engine's algorithm scans the search index to find the best webpages, and then it ranks those pages on a variety of factors, including relevance, popularity, authority, and more. These ranked results are then presented to the user with the most relevant and authoritative results first. The results page also often includes other content, like FAQs, sponsored ads, plus more data, like images, video, news, articles, maps, and more depending on the search engine, and the nature of the query. So here you can see my search engine query, "Parts of a grant proposal." It returns all of these useful resources. Most of them are webpages that I can access to read further about my query. This type of results return probably looks very familiar to you. Search engines have long been an easy, effective free way for you to find all sorts of information. They offer comprehensive, wide ranging access to countless topics, and because they prioritize the most relevant and accurate results, it's a great time saver. Many search engines also have advanced search options, like filters, suggestions, and search shortcuts, which let you refine your search for even better results. Now, that we've reviewed how a search engine works, let's explore how a reasoning engine works, and then we'll be able to discuss them in relation to one another.

How a reasoning engine works

How a reasoning engine works
Selecting transcript lines in this section will navigate to timestamp in the video
A reasoning engine is a system that uses logic and inference methods to draw conclusions, make decisions, summarize information, or solve problems based on available data and knowledge. - So a reasoning engine, it's in the title, right? You assume that it's making some rational thinking. It's reasoning from the inquiry that you're giving it as compared to a search engine, which is more likely just to give you back a response to the direct information that you queried. - There are many different types of reasoning engines, but here I'd like to talk about generative AI models like ChatGPT and Bing Chat. These systems process and understand human language. So when a user enters a query, the reasoning engine can provide a relevant and informative response using human-like speech. It's kind of like having a digital friend who can answer your questions, engage in discussions, and help you with various tasks. Let's talk just a bit about how this works. The GPT large language model that powers these reasoning engines was pre-trained on a massive amount of training data. This includes text and code from the internet, from books, articles, webpages, Wikipedia, and more. During this pre-training phase, the model learns language patterns, grammar, syntax, and factual knowledge. The model then uses probability to predict the next words in response to a given prompt. In the initial training phases, human supervisors oversee the process, guiding the model toward accurate responses and contributing to its knowledge development. During later training stages, the AI model is asked to generate multiple responses to a prompt, and then human supervisors evaluate and rank these responses from best to worst. Gradually, the model improves its understanding of prompts and increases the probability of generating accurate, comprehensive, human-like responses. Once all of that happens, you have your smart digital friend that you can chat with. - In most cases, these reasoning engines are pre-trained on a bunch of other similar-like reasoning, where you'll see the difference in the kind of response you get from a reasoning engine as opposed to a traditional search engine because it'll be much more conversational, but also much more elaborate in its connection to what you're asking for and the different types of responses that it can create. - Take a look at a few different types of results I can get from a reasoning engine. Instead of asking what the various parts of a grant proposal are, I can ask Bing Chat to actually help me ideate and research and write various parts of a grant proposal in a very customized way. Reasoning engines aren't only used to generate text. They can generate other media types as well, including images. Text-to-image generative AI programs, where neural networks learn the relationship between words and a massive database of images, let you enter a text-based query to produce an image-based result. These models can create complex imagery while also responding to a wide variety of concepts, tones, and styles. Take a look at these images that I can create using the Bing Image Creator when I enter the prompt, a young boy walking a dog in a forest under a full moon, digital art style, bold colors. - You can use these tools to create incredibly beautiful or realistic imagery or imagery that is, you know, perfectly aligned with what you had envisioned. I think that that's a skill that artists can use to say, okay, if you want to use these tools, I have deep skills in being able to write prompts in a way that will actually get to that end product. So I actually think it might be a new tool that artists can use to express themselves. - So whether it's text, image, or other media, reasoning engines have burst onto the scene and attracted a lot of attention and excitement with the possibilities of what users can create. Let's continue to talk about how you can use reasoning engines both independently and in conjunction with search engines.

Comparing search engines with reasoning engines

Comparing search engines with reasoning engines
Selecting transcript lines in this section will navigate to timestamp in the video
Let's talk about some of the pros and cons of reasoning engines versus search engines. We know that when you enter a query into a traditional search engine like Google or Bing, it searches through its vast index of web content, and then provides you with a ranked list of webpages that contain relevant information. So it's a good resource to use when you'd like to read further about a subject across a collection of different sources, but not necessarily when you want to ask deeper questions. That's because a search engine doesn't truly understand your query. It just matches your keywords to relevant results. - Traditional search is much more linear, right? There's a linear request being asked against a data set to bubble up anything that's remotely similar. There's often some level of ranking that's done to determine, you know, how closely to your question did a answer get, and that's how that rank of, you know, what's on page one of a search engine versus page two. But very rarely does it actually think about what did you mean by what you asked? And this is often a frustration with us as users thinking about a search engine. I'm trying to find something and I'll often be like, "Well, no, that's not I meant to say," and I'll have to adjust it. - A reasoning engine, on the other hand, is designed to understand and interpret human language, which allows it to engage in actual conversations with you. So instead of simply retrieving webpages for you to read and extract information, it provides you with direct relevant responses and can also maintain context and understand the intent behind your questions. - Reasoning engines take in a completely different perspective when you're asking it questions. So what we've seen in the evolution of something like a GPT, for example, is that when we start to ask a questions rather than it linearly looking for a direct answer and ranking everything that's similar to our response, it now is taking in the context of that question related to each other, right? So maybe topics, locations, places. - Sometimes, especially when you're first starting out, it may not be immediately obvious which one to use. But often, just by going through the motions of asking your question, you'll get a sense of which direction you should go. - Honestly, the first thing I will do, sometimes, just out of habit, I will go into a typical search engine, but almost immediately, as soon as I start typing, I'm like, "Maybe I should use a different engine for that." The main reason for doing that though is that often when I'm asking a question, I'm very rarely asking a pointed, very specific, factual answer. I'm usually asking it for something that is more related to help me with this idea. Help me create something, help me ideate. And if I do it in a traditional search engine, I will get examples that then have another step that I have to go into and actually find my answer, which creates more work. So with a reasoning engine, a lot of that is surfaced in a way that I don't have those extra steps, makes that process much faster, but also just much more human-friendly to the creative process. - One thing to watch out for, however, is that even though it sometimes seems that reasoning engines are an all knowing source of truth, that's not necessarily always the case. - One caveat that I think is really important about these reasoning engines is that when they do produce content, it sounds incredibly convincing, that that system has perfectly understood your query, and it has given you back valid results. Consider that maybe, it's not actually giving you the correct result. So everything it gives you at this point, you need to treat with a little bit of caution. Now over time, those reasoning engines might actually get better and better and better, and they will be correct. - One of the disadvantages of reasoning engines that I tend to think about often and we've heard this term like hallucination, right? Like can it be wrong? Now, I often also then think about the amount of things that have been wrong on a typical search engine, right? And that I think the difference though is because the reasoning engine is formatting its response and that response is much more kind of human and natural language-driven. That natural language tends to make us feel like it may be more accurate than a list of resources that we have to kind of determine the accuracy on our own. - So just be mindful of this. Whether you use a search engine, which provides that list of ranked web content, or a reasoning engine, which provides that human-like contextual response, it's up to you to dig deeper to verify and validate your results.

What is the future of online search?

What is the future of online search?
Selecting transcript lines in this section will navigate to timestamp in the video
We're currently at a stage where most people have used search engines extensively for years and now many people are starting to use reasoning engines for the first time. This means it's a very interesting and experimental time in the world of online search, and we're all learning how to navigate these new waters as we determine which is the best tool to use. - We want to understand why we might use one versus another or, in my case, I like to use them both and see what kind of answers they come up with together. And I see now technology is really evolving to combine these models so that I don't have to do that work. Our future search engines are going to work synergistically and remove that aggregation piece from the user, right? Like, what if I'm wanting to write a new medical journal article, or what if I want to write a new business leadership or thought leadership article and I go and do a search? Today, if I go to a standard, you know, typical search engine, I will get other articles like that. It's not going to generate me anything. But imagine a world where I ask that question, and it not only gives me an outline to start, but it also provides me really great examples that have been used. I think that's an interesting way to think about the future, right, of how we're going to combine these two technologies together. - I do think that there will be this integration happening more and more where your traditional search engine has this reasoning agent part of it as well. And I think that's in part because it helps to do a little bit of the work that you would do when you go to a search engine. If you're querying something and you want to know, what's the state of X, you know, what's happening in this space, you have to go and read a few links, start to synthesize it yourself. So I actually think that these tools might be used in the future to help synthesize that content for you and just make it a lot quicker for you to get to, you know, the understanding that you're trying to reach. - Already there are reasoning engines like Bing Chat that are combining both functionalities. As you can see here, when I enter a query, I not only get a reasoned response, but I also get a list of sources that informed each part of my response. These sources are also presented as links where I can read further and continue my research on the subject. Harnessing this type of synergy between search engine and reasoning engine will continue to be an effective way to help verify and validate the results of online search, and also to dig deeper and explore related sources. And the ways in which they're used together will continue to change and evolve, making online search more thoughtful, robust, and comprehensive.

Harness the power of prompt engineering

Harness the power of prompt engineering
Selecting transcript lines in this section will navigate to timestamp in the video
One of the most important skills to develop when working with reasoning engines is in the area of prompt engineering. There's truly an art to the way you craft your prompts when you query reasoning engines. - Prompt engineering is the act of asking questions of one of these foundation models. So the first time you go in and you enter in a prompt, you can consider yourself a prompt engineer. But what accelerates that capability and your ability to do a better job is actually telling the model, not just asking it a question, but then also giving it examples of the question and the answer you would like to see. We often call this one shot or few shot learning and what it means is how much instruction are you giving when you're asking it a question. So the prompt itself is really quite simple. It's just the question you ask. But the instructions that you provide to guide the answer, that is where true prompt engineering starts to kick in. - There is this incredible skill set that you have to develop to be able to harness these tools, and it really, really is relying on the development of appropriate prompts. You will get out only as good of a prompt that you put in. So it's almost like the great painters of the past were those who could wield the paintbrush. I think the greatest artist maybe in the future are those who can wield the prompt. Part of that tinkering in knowing, okay, well this is the image it gave to me from this prompt. I really want to move it in this way. How do I manipulate the wording of the prompt to get it to give me that? And that is such an amazing skill set that people are developing. - Now, let me tell you, this just takes practice. When I first started to enter prompts, I asked questions like I would ask a search engine and I didn't get great results. It wasn't until I learned how to hone my language that I started getting the results I was after. And as I continued honing my prompt language, I started getting results that I didn't even know were possible. Take a look at the difference between this prompt, write a social media post about International Women's Day and this one, write a cross section of social media posts for Facebook, LinkedIn, and Twitter for International Women's Day. My business is a boutique video editing company and I want to celebrate female film editors. Keep the tone celebratory, casual, fun, and optimistic. And then, how about this prompt? An image of a red car, and this one, an image of a red car in a busy downtown area approaching a traffic jam. The car should be the only color in the image. The rest should be black and white. In each of these examples, additional details regarding content, context, tone, and aesthetic, give me better results. Again, as you learn how your prompts impact the output you create, it's important to analyze the results for various factors, not just related to the content itself, but also to various quality standards. And if the results aren't up to your expectations, continue to hone the prompt. - Look at the results and evaluate are these results accurate? Are they discriminatory in any way? Are they only presenting one political viewpoint? And then I think it's continuing to iterate. The power of these systems is continuing to iterate on those prompts. It's not just one prompt and you're done. It's start from a prompt, see where it gets you and then keep regenerating and refining the prompt to get at a more accurate or better result. - [Instructor] Here are some resources for you to start learning the art of prompt engineering. Don't be shy, just start prompting, analyzing, and tweaking and I promise you'll improve the way you query over time to generate some awesome results.

Thoughtful search strategies and approaches

Thoughtful search strategies and approaches
Selecting transcript lines in this section will navigate to timestamp in the video
We talked about the importance of developing good prompt engineering skills. Here, I'd like to expand on that plus offer some additional tips for thoughtful reasoning engine search strategies. First, be specific. The more specific the prompt, the more customized and nuanced the generated results will be. Provide context. Reasoning engines do well when you include context within your query which often includes providing an example of the type of answer that you're looking for. - Rather than having to infuse it with knowledge, it already knows what it knows. Instead, we need to give it examples of what right looks like and how we can answer the question that we've asked in the very specific way that we want it answered. And your ability to provide that instruction happens in the prompt as well. - Break things down. If you've got a complex query, try breaking it down into smaller parts. This can make it easier to get a thorough non-confusing answer. Use clear language. Reasoning engines are designed to synthesize human speech so it tends to do a fine job understanding colloquial language, jargon or slang. But if you want the best results, write your prompt as clearly as possible with proper grammar. One thing though, if you're creating a prompt for a specific industry or use case, using very clear examples of industry language is the way to go. - One of the most interesting things about this role is it's very domain specific. Someone who's building the prompt has to know the problem they're trying to solve. It's less about the technology and more about the business problem or the challenge you have. I want to be an animator, so I'm like, I can build animation. But in order for me to get like a GPT model, for example to do an an animation, I actually have to know all the lingo and you know all the jargon of animators. In order to do that work, I need to be an animator to use that tool. - Experiment. Try out all sorts of ways to write your prompts. Often slightly changing your query leads to different responses and different levels of creativity. And the neat thing is this type of experimentation isn't just making you learn it's also making the reasoning engine learn. - It's building over time and learning from the inputs. And it's more of a sort of a dance between human and machine because it's always, you're iterating the prompt. You're trying to push it into these new directions to be able to be more creative and in alignment with what you envision. And I think over time, as it continues to learn from this it might actually get more creative itself. - Every time we use the model, the outcome changes the ground truth shifts even just a little bit. So we're going to have to make sure that we're continuously training and monitoring the output of that model over time. - Along those lines, here's a list of some creative ideas for how you can expand your prompt engineering skills. Role play scenarios. Frame your query as a role play scenario where the reasoning engine takes on the persona of someone relevant to your query. For example, imagine you're the product manager for a brand new smartphone company. What are 10 potential innovative features that could be added within the next five years? Analogies. In your prompt, explain concepts using an analogy to better understand complex ideas. For example explain quantum mechanics using a sports analogy. Debate style questions. Ask for arguments for and against a particular topic to get a more comprehensive understanding of different perspectives. For example, present arguments for and against the implementation of universal basic income. Creative exercises. Ask the model to brainstorm, ideate or write fictional scenarios related to your query. For example, I'm writing a story about traveling to a parallel universe. Help me brainstorm some unique laws of nature that this parallel universe might have. And finally, during all of the excitement of finding new ways to generate content, don't be afraid to say that you've used the tool. - I think there's also a level of transparency another responsible AI principle that should encourage us as organizations, as users, to not be shy about saying when we've used GPT, many times I find people are almost like I used GPT for that. Almost like it's a secret. And I think part of it is normalizing the use of it so that you're transparent about what's being generated and what isn't. - There's been a lot of work especially in the academic setting right now about appropriate guardrails for students in using this. And I think that that also needs to be applied in corporate settings. If you are going to be using these tools to aid in your creative process or in writing or creating imagery, it must be transparent that you did use them. - Now go forth and engage in those thoughtful search strategies. The possibilities are limitless.