AI and Its Impact on Cybersecurity

Part 4: Search Engines That Can Read Your Mind

In Part 3, we learned that your spam filter has been using AI to protect you for years. Now let me ask you a question. When was the last time you searched for something on Google and couldn't find it? Probably not recently. If you are old enough to remember the early 2000s, or if you are like me, the 1990s, searching meant typing exact keywords and hoping for the best. What changed? Just like spam filters evolved from keyword blocking to sophisticated AI, search engines made the very same journey. Today's Google doesn't just match keywords, it understands context, interprets vague questions, and even corrects your typos before you finish typing. Notice how as you type it provides a list of searches based on the words you are typing? That's the power of AI learning patterns from billions of searches. In part 4, we look at how search engines transformed from simple word counters into systems that seem to know what you're thinking.

Case Study: The Evolution of Google

We will use Google, one of the most influential, early and most recognizable search engines, for an example of how AI has been used over the years. Google, founded in 1998, didn't suddenly implement artificial intelligence into its search engine. Just like the development of AI, the incorporation of this technology was a gradual evolution since its founding. Early search engines like Google counted keywords to determine the rank or importance of a website. The higher the keyword count, the higher the ranking and the higher the page would appear in search results. You can see the problem with this method. Web page designers just jam-packed websites with keywords so the page would be ranked higher. If you were selling garden gnomes, you would fill your website with as many keywords related to garden gnomes as possible. This really didn't make for impactful websites, just repetition, and search results that could return garbage.

Larry Page, one of Google's founders, thought how do academic papers establish credibility? These papers achieve credibility by how many times the paper is cited. If others are citing your work, then your work must be more important. Larry Page's idea was to apply this methodology to the internet: webpage links = academic citations. The more links to your website, the more important your website must be and therefore it is ranked higher. For example, you are looking for a recipe for chocolate cake. Recipe A has five links to it. Recipe B has 300 links to it. Recipe B must be much better. Google called this PageRank. But it still wasn't good enough.

The reason this was still a problem was that there was no way to evaluate the value of a link, or a recommendation. In the example above, a recommendation from a master baker would be very valuable. A recommendation for chocolate cake by an 8-year-old boy, not so much. So, Google started weighting links based on their value. How do you determine the value of the page? By how many valuable pages link to it. The master chef is going to have more links to it by other master

chefs, which gives it more weight or a higher rank, whereas the 8-year-old will have links from other eight-year-olds, who really do not know much about baking, so that page is given a lower rank.

What made PageRank work so well is that it was not determined by Google, but by the users. To boost your website, you need links from other high-quality websites. You couldn't just create a bunch of websites to link to your site; those sites would be considered low quality sites and have low rankings so they would not help. The importance of a page depends on the importance of the pages linking to it, which depends on the importance of the pages linking to them. When they tested PageRank they determined that it matched what human experts thought was important!

From Algorithm to AI: The Evolution

PageRank was revolutionary, but here's an important distinction: it wasn't AI. PageRank was an *algorithm*, a fixed set of rules. It counted links, weighted them, and ranked pages the same way every time, like following a recipe. No learning, no adapting, just following instructions.
The evolution to AI happened when Google realized algorithms weren't enough. People don't search like computers think. Someone searching "what's that movie where the boat sinks and there is some controversy about a floating door" is looking for "Titanic," but an algorithm only sees keywords. It doesn't *understand*.

Rabbit Hole! If you think algorithms are something new, take some time to look up a man from modern day Iran by the name Al-Khwarizmi. He lived from around 780-850 AD. His work in mathematics was instrumental in the algorithms we use today. As a matter of fact, algorithms are named after him! We will take a deeper look into algorithms in part 5.

PageRank was good, but not good enough. In 2015, Google introduced RankBrain, their first real AI system for search. Unlike PageRank's fixed rules, RankBrain learned from millions of searches to understand what people really wanted, even when they couldn't quite express it. This was pattern recognition in action, what we discussed in Part 2 of this series.

If you search for "gray long-eared animal that hops," RankBrain recognizes from patterns in previous searches that you probably mean rabbits, even though you never said the word. It learned that searches with those characteristics typically lead to rabbit-related results. That's the difference: algorithms follow rules, AI learns patterns and then they predict.

This evolution continued in late 2018 with the introduction of BERT (Bidirectional Encoder Representations from Transformers), which brought those Transformer models from Part 1 into search, and then in 2021 with MUM (Multitask Unified Model), which understands multiple types of information at once, like text, images and videos. MUM's advantage is that it can provide more comprehensive and relevant search results on the first try, reducing the number of additional searches to get to the information you want.

Each advancement moved further from fixed algorithms into true AI territory; systems that learn, adapt, and improve. This same evolution happened everywhere online during the 2010s and 2020s. Simple algorithms transformed into AI systems that recognize patterns and make predictions. What does that mean for today? What does the modern progression of algorithm to AI look like for users as they search the internet?

Modern Search AI in Action

In our opening paragraph we already mentioned autocomplete searches, when you start typing a search phrase and Google predicts what you might be searching for by providing a list of probable searches. Sometimes these can be quite funny and unexpected! These suggested searches are the result of your own search history, your location, and patterns from millions of similar users. What are other examples of where AI is used in your search process?

Alongside autocomplete there is autocorrect. You see this when misspelling a word like "restaurant". After you type the word incorrectly, AI will recognize the error and provide a suggestion to alert you to the error. This is also used on our smartphone texting apps. This can be turned off leading to some very entertaining text strings!

Most search engines have now fully integrated their own AI into its functionality. Most Google search results will now result in a fully AI generated response to your search at the top of the page called Search Assist that uses their AI model Gemini. If you phrase your search in the form of a question, Gemini will return an answer! Some search engines will also provide a "People Also Ask" section in results showing potential follow up questions.

All of this happens in milliseconds for every search. This is part of the pattern recognition we discussed in part 2 of this series at much larger and faster scale. AI analyzes your question then "understands" the context of your question, compares it to billions of other searches and predicts what you want and returns the results.

The Power of Pattern Recognition on Steroids

Back in part 2 we used the four "S"s to describe how AI is pattern recognition on steroids. Here is a reminder of what we talked about but related to internet searches.

1. Scale: Google processes over 8 billion searches a day. Yes, over 8 billion! It learns from every one of those searches, refining its pattern recognition, getting better with each search analyzed. You can see how over time the accuracy of the predictions would increase.
2. Scope: Google analyzes hundreds of factors simultaneously; location, search history, time of day, device type and many, many more from similar users for each search.
3. Speed: This is all possible due to the advancements in processing chips, memory and hardware.
4. Subtlety: Bundle all of this together and AI is able to see subtleties in words or phrases based on their context in the entire search phrase, like "river bank" and "mortgage bank".

These advancements and innovations have moved us from simple pattern recognition based on rigid rule following to a fluid analytical prediction model that is able to continually build on its accuracy making better predictions. Still based on patterns, still based on word recognition, still based on following rules, but with the ability to use context by seeing the whole picture. By "paying attention" as those Google engineers wrote back in 2018.

Is the Evolution of AI Complete?

The question now is, have we come as far as we can possibly go? Some say yes, some say no, that is the current debate and outside the scope of our discussion. What we can say is that in a very short period, concepts that have been around for a very long time have been developed into very sophisticated programs that are able to do things that are quite incredible. From keyword counting to PageRank to AI models able to provide you with a weekly meal plan, grocery list, and estimated weekly grocery budget. There is no doubt that AI is here and it is being used virtually everywhere even when you cannot see it.

In part 5 we will look at how AI uses algorithms behind the scenes and can affect how much you pay for something! Part 5 will be the last look at what AI is under the hood. We have looked at examples of how it is being used for the benefit of users but after part 5 we will start looking at how AI is being used to harm you and steal from you. Know Phishing was started with the goal of informing everyday tech users how to stay safe and keep your data protected in an online world. All without using overly tech heavy jargon, but by explaining in the simplest, easiest to understand way possible. In the meantime, head over to KnowPhishing.com and sign up for our newsletter for weekly information, tech tips and alerts on staying safe from the bad use of AI.
Thank you for your “attention”!