Introducing: Delphi AI

View of Delphi with Sacrificial Procession by Claude Lorrain

Daniel Miles
Jul 24, 2023

Impatient? Just try it

Cyber security is actually pretty hard. It's even harder to get it "right".

We live in one of the most powerful and enabled generations of all of history. The birth of the internet era has enabled *literally anyone* with a burning enough desire to solve problems at a scale previously thought impossible.

That said building digital products and services isn't a complete walk in the park, and the internet can be a dangerous place if not treated with the respect it needs. The rise of cyber crime has led to market size of $7 Trillion by some estimates, larger than some of the worlds biggest economies.

> *Building digital products is great, doing it securely is*

We know that not everyone is a cyber security expert, we also know that a lot of the terms and language is incredibly nuanced and vague, even confusing the experts themselves.

Getting answers to questions such as *am I doing this right*, *what the hell is "compliance"* and *that alert looks scary what should I do?* is never as easy as we'd like it to be.

This is where Delphi comes in.

The What

Delphi builds on top of our work with other AI models, specifically Large Language Models.

Delphi itself is a fine-tuned LLM using a custom dataset of cyber security information, alerts, and reporting that we've been compiling for several years now.

We've then built a very simple chat wrapper around the produced model that you can ask questions to, very similar to how other conversational AIs work, but specifically designed to help you get answers and next-step actions for cyber questions or problems you're facing.

Today, it's just a simple chat conversation. We started with just this as we wanted to get a feel and understanding for the kinds of questions people have about cyber and how they would ask them.

The upcoming future roadmap however, is incredibly feature rich, including;

- Directly plugging into your current cyber security tooling to automatically collect, explain, and advise about alerts or notifications
- Allowing you to upload files or links that you might have opened but suspect might be malicious, and get instant advice and feedback about it
- Pasting whole rows of vulnerabilities and having it understand and prioritise which ones you should start with.

Sounds good enough? Try it out here

The How

As we mentioned earlier, Delphi is an AI model specifically tuned to understand cyber security better, but it doesn't stop there.

Whenever a question is given to Delphi, it first performs a lookup of relevant and contextual information that can better inform its decision making from both our massive and constantly evolving dataset of cyber information, and the contextual understanding of who you are and what you've previously asked it the more you use it.

All of this is then fed into a chat completion system that has been pre-prompted with the relevant context and requirements. We're continuously evolving what these prompts look like to ensure we're getting the best results possible.

Our product thesis for this being that the chat agent needs to both;

- Perfectly tailor the response to the user that is currently asking the questions, be it a cyber security professional, or someones grandmother that has unknowingly opened a file they shouldn't have
- Provides concrete next steps that you can action with confident authority.

That last part is a lot harder. Getting language modals to be commital about their answers after performing sufficient reasoning about them can be hard, but we're confident in our approach

In any case, we're excited to see what people ask Delphi and see how it continues to evolve as a useful tool for both professionals and not over time.

I've heard enough, show me!

← Previous Next →