ChatGPT Created Threat Model

# ChatGPT

ChatGPT-3 is a chatbot developed by OpenAI, a US-based artificial intelligence research lab. Many consumers are familiar with chatbots, but less so with the underlying technology. ChatGPT-3 uses a generative pre-trained transformer (GPT) to generate text that is largely identical to human conversation. GPT is part of the broader family of large language models, which are AI models that understand and can generate text. Large language models are compelling because of their flexibility. With minimal training, large language models can complete sentences, translate a foreign language, summarize information from multiple sources, and generate content ranging from technical answers to humorous vignettes. Enormous amounts of data from a range of sources are used in the training process. ChatGPT-3 ingested content from books, academic papers, and the entirety of Wikipedia. According to OpenAI, ChatGPT-3 was trained on over 45 terabytes of data. The system iteratively and without supervision learned to predict the next word in a sequence of text. This unsupervised learning required a network of distributed computers, which allowed for a faster learning process than could have been accomplished using a single computer. The combination of a vast amount of ingested data and massive computing power resulted in a model with 175 billion parameters, making it one of the largest language models ever created. Parameters in large language models are essentially the values that the model is solving.

# Threat Modelling

The fundamental basis of threat modeling is identifying, communicating and managing security weaknesses. This is achieved by understanding potential threats and attacks the system must resist and the corresponding countermeasures (controls) for those threats. Fixing in production vs Secure by design. The key principle underpinning threat modeling is “secure design” which means in practice addressing design will take place from the inception of the project at the design phase and continue throughout the development life cycle.

Threat modeling should begin when the functions and components of a product are known and the architectural design is being formulated. In this way threat modeling drives security requirements from the outset and has the dual method. This also ensures building and weaving security into the process rather than bolting it on at the end.

The process of threat modeling usually starts with a diagram which describes and maps the architecture, components, trust zones and authentication flows. It is very useful to incorporate data flows within the diagram: Data flows show how information enters and leaves the system, what changes the information and where information is stored. The purpose of a data flow diagram (DFD) is to show the scope and boundaries of a system as a whole.

Within the diagram potential weaknesses and attack points are identified. This threat enumeration activity may be promoted and facilitated by referring to attack trees (MITRE or CAPEC) or standards (PCI DSS, GDPR, SO, NIST) or guides such as OWASP’s Mobile Application (MASVS). Also, the Common Weakness Enumeration (CWE) and Common Vulnerabilities and Exposures (CVE) are useful here. Microsoft’s STRIDE Methodology is a popular tool for thinking about security threats within six categories; namely: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service and Elevation of Privilege.

There is a wide range of threat modelling methodologies that companies can make use of, as each is a unique approach and provides varied benefits. Among these, the most common are STRIDE, OCTAVE, TRIKE AND PASTA.

Integration of ChatGPT to Threat Modelling Process

