AI & F5

Giedrius Zybertas
BDM, B4, BAKOTECH
Giedrius.Zybertas@Bakotech.com
+37061460838

Agenda

AI for business

AI for F5

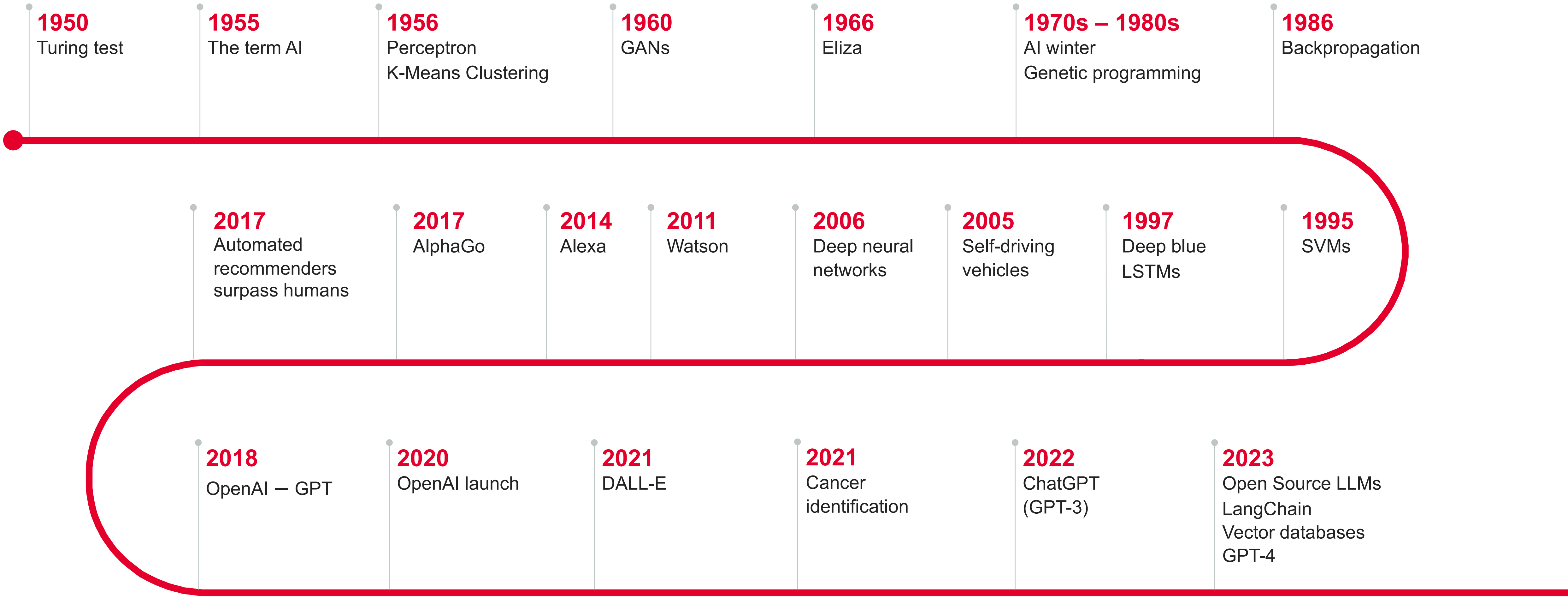
F5 for AI

Key takeaways

AI for business

AI is not new

History of AI



1950
Turing test

1955
The term AI

1956
Perceptron
K-Means Clustering

1960
GANs

1966
Eliza

1970s – 1980s
AI winter
Genetic programming

1986
Backpropagation

1995
SVMs

1997
Deep blue
LSTMs

2005
Self-driving vehicles

2006
Deep neural networks

2011
Watson

2014
Alexa

2017
AlphaGo

2017
Automated recommenders surpass humans

2018
OpenAI — GPT

2020
OpenAI launch

2021
DALL-E

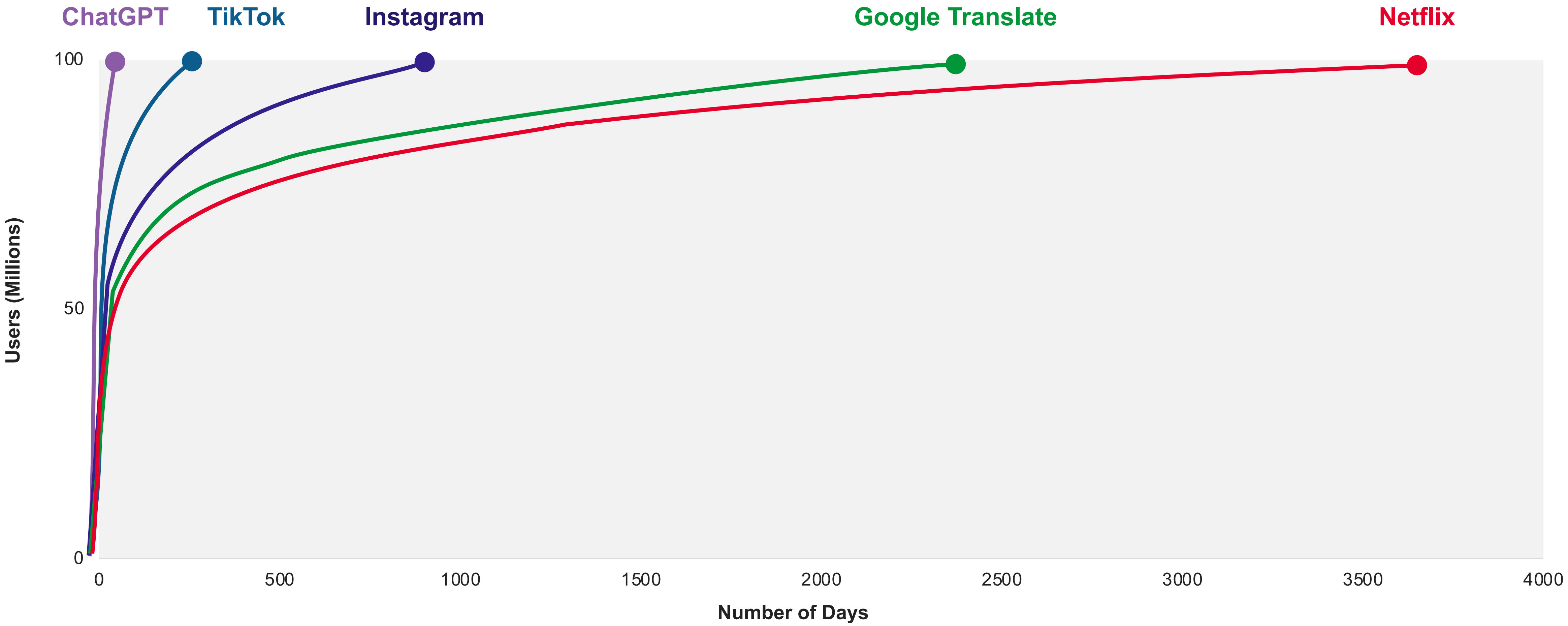
2021
Cancer identification

2022
ChatGPT
(GPT-3)

2023
Open Source LLMs
LangChain
Vector databases
GPT-4

The pace of adoption of Generative AI has been astounding

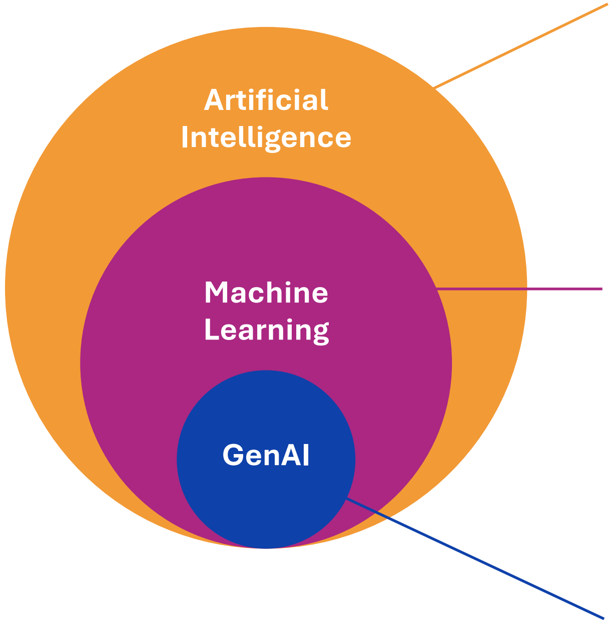
Time it took companies to reach 100 million users:



Sources: Global X ETFs with info derived from: BBC News. (2018, Jan 23). Netflix's history: From DVD rentals to streaming success; Cerullo, M. (2023, Feb 1). ChatGPT user base is growing faster than TikTok. CBS News.

Generative AI and Large Language Models (LLM)

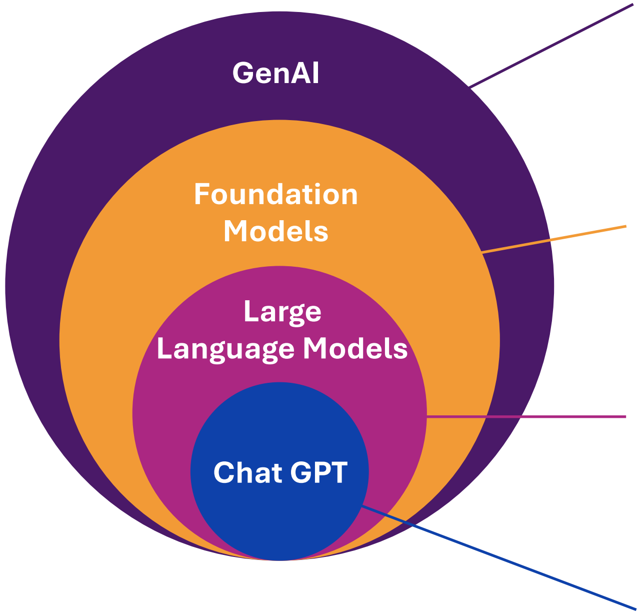
Setting the context



Technology which emulates human or beyond-human intelligence

A form of AI which uses data to train models to <u>make predictions</u> or <u>produce insights</u>

A form of machine learning which uses large quantities of data to train models to <u>produce content</u>



A form of machine learning which uses large quantities of data to train models to produce content

A general purpose model trained on large quantities of data

A model trained on large quantities of text to produce plausible content

A publicly available service which provides a chat interface to an LLM

Generative AI has democratised technology

Anyone, technical and non-technical, can unlock its power

1,000+

Plugins available with **GPT-4**

- Create an app when you've never coded before
- Appeal an insurance denial
- Write Excel formulas
- Design a personal shopping assistant
- Build new games
- Write marketing blogs

Source: OpenAI

Adoption of AI in business is starting to ramp up

75%

of organisations consider AI a core business focus

AI models are

Hybrid

with 40% of orgs deploying AI models on premises and 65% in public clouds

AI budgets expected to grow

94%

from 2024 to 2026

Most businesses consider

Security and Compliance

as one of the top challenges that complicate deployment of AI models and applications

Source: F5 2024 State of Application Strategy Report

For cyber security, AI is a double-edged sword

Creates new attack surfaces and new attack capabilities

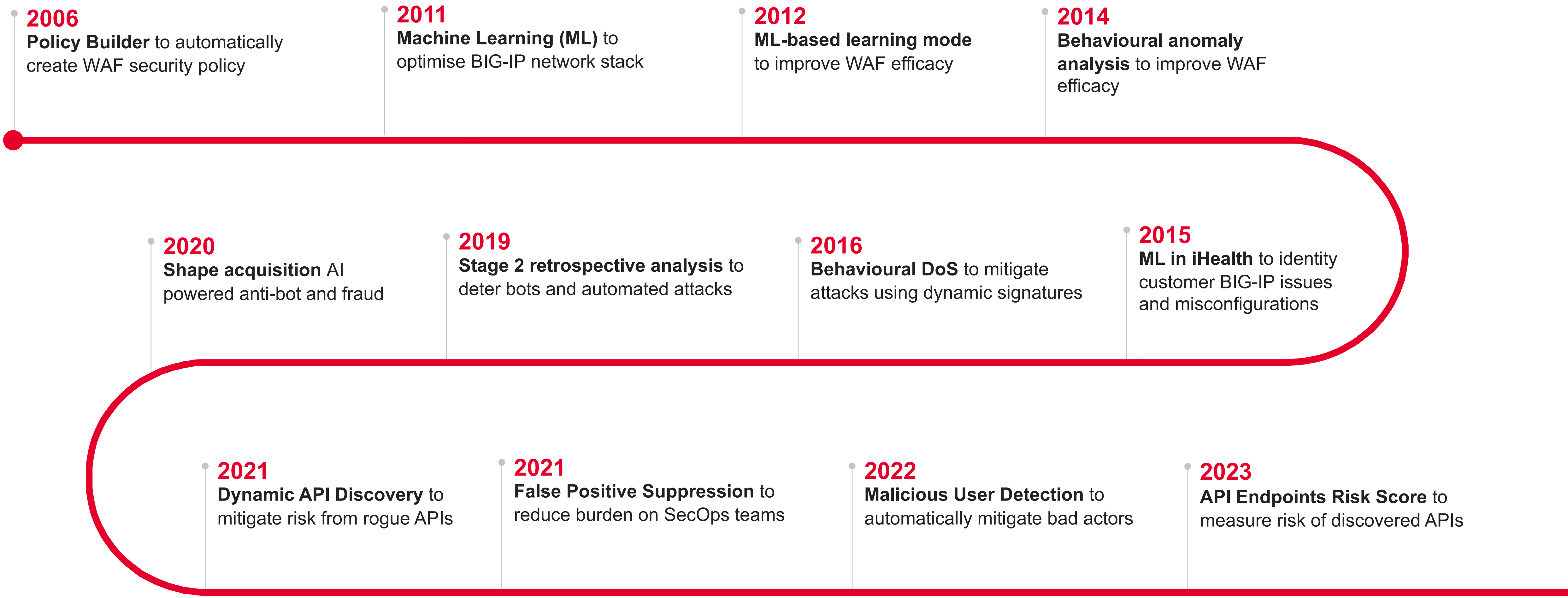
Moving at alarming speed,
no reservations

Enhances cyber security

Exercising caution with deferment to human expertise

AI for F5

AI is not new to F5



2006
Policy Builder to automatically create WAF security policy

2011
Machine Learning (ML) to optimise BIG-IP network stack

2012
ML-based learning mode to improve WAF efficacy

2014
Behavioural anomaly analysis to improve WAF efficacy

2015
ML in iHealth to identity customer BIG-IP issues and misconfigurations

2016
Behavioural DoS to mitigate attacks using dynamic signatures

2019
Stage 2 retrospective analysis to deter bots and automated attacks

2020
Shape acquisition AI powered anti-bot and fraud

2021
Dynamic API Discovery to mitigate risk from rogue APIs

2021
False Positive Suppression to reduce burden on SecOps teams

2022
Malicious User Detection to automatically mitigate bad actors

2023
API Endpoints Risk Score to measure risk of discovered APIs

F5 AI Data Fabric – Roadmap to AIOps

Leverage Generative AI to deliver insights and ease of use across our portfolio

1

Train and refine ML detection models

Use LLMs to build highly accurate classifiers in minutes, then generate signatures.

“That’s an attack. Train and deploy a model to detect events that look like it.”

2

Dynamic data visualisation

Use LLMs to turn natural language prompts into API queries of the F5 data platform.

“How many times did that attack happen in the last week? Graph that for me.”

3

Virtual assistant through a chatbot

Easily converse with knowledgebase articles and receive recommendations.

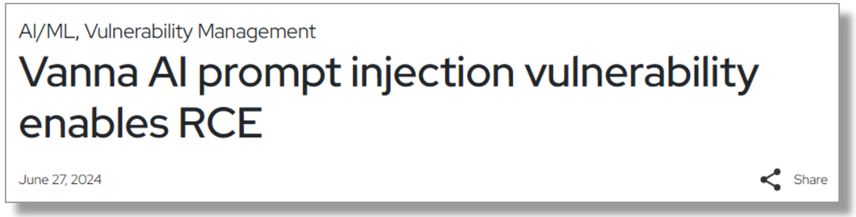
“How can I configure blocking actions on that type of attack?”

F5 for AI

Generative AI in enterprise – With rewards come new risks

Secure enterprise AI apps

- Eliminate risk of prompt injection, data leaks, and harmful LLM responses from GenAI apps



AI/ML, Vulnerability Management

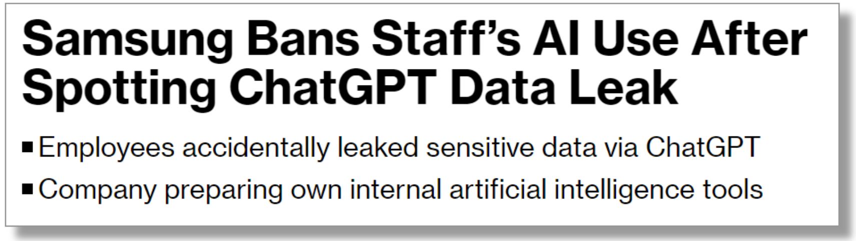
Vanna AI prompt injection vulnerability enables RCE

June 27, 2024

Share

Protect employee usage of GenAI apps

- Enable employees to adopt GenAI tools without worrying about data privacy and regulatory risks

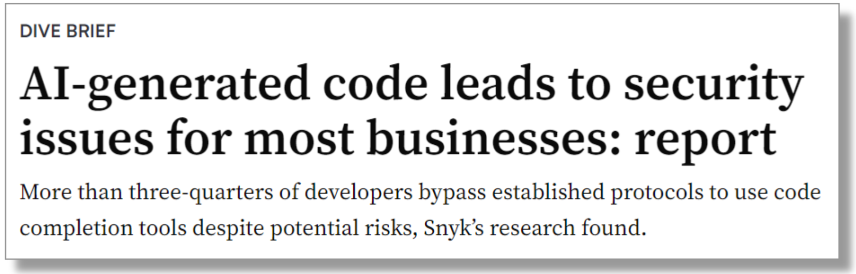


Samsung Bans Staff's AI Use After Spotting ChatGPT Data Leak

- ▪ Employees accidentally leaked sensitive data via ChatGPT
- ▪ Company preparing own internal artificial intelligence tools

Defend developers using AI code assistants

- Adopt AI-based code assistants like GitHub Copilot without worrying about secrets exfiltration



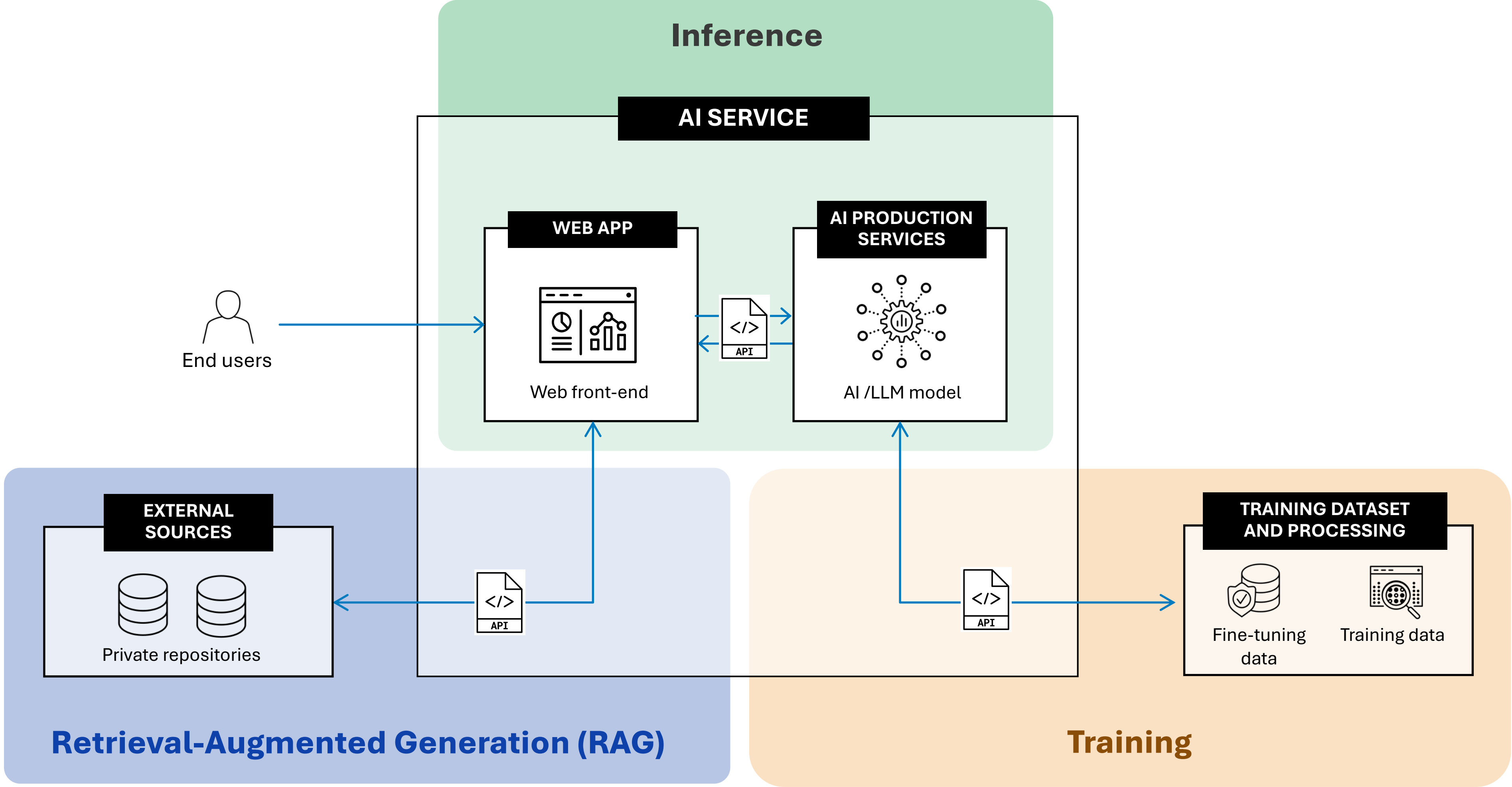
DIVE BRIEF

AI-generated code leads to security issues for most businesses: report

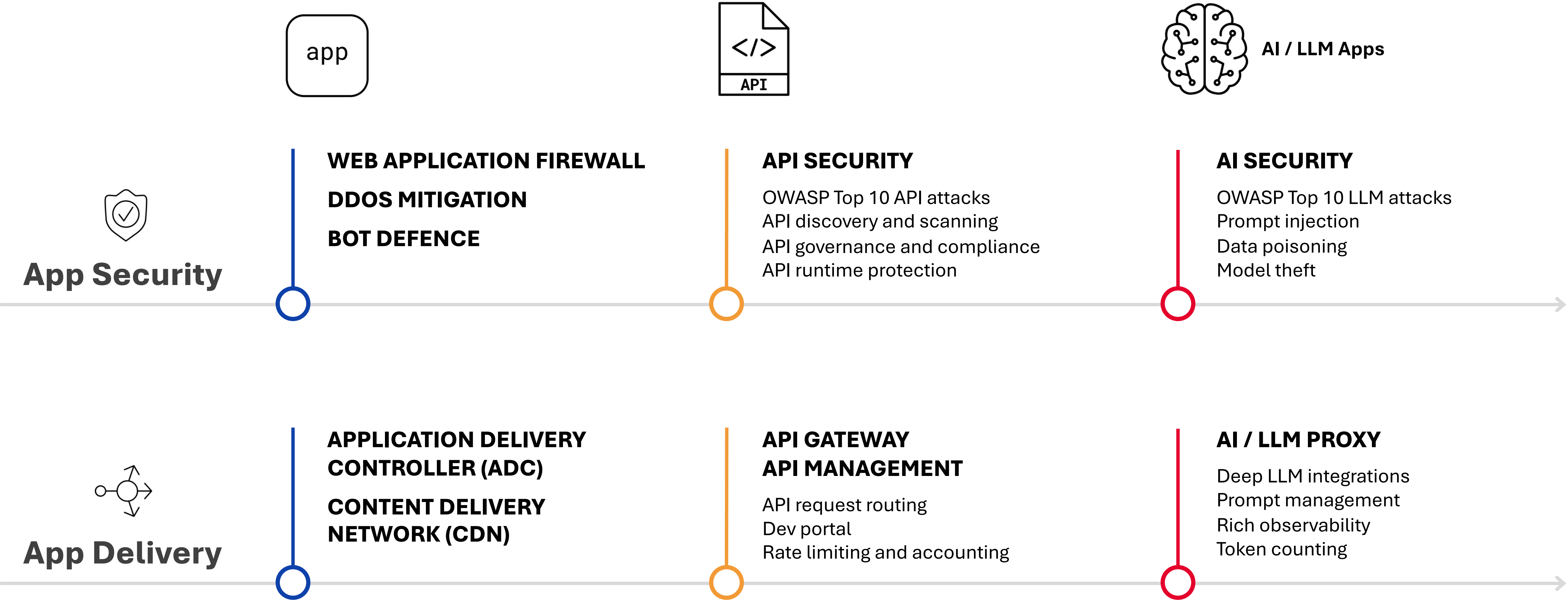
More than three-quarters of developers bypass established protocols to use code completion tools despite potential risks, Snyk's research found.

Training, Inference, and Retrieval-Augmented Generation (RAG)

LLM application architecture



From web applications, to APIs, to AI



app

API

AI / LLM Apps

App Security

WEB APPLICATION FIREWALL

DDOS MITIGATION

BOT DEFENCE

API SECURITY

OWASP Top 10 API attacks
API discovery and scanning
API governance and compliance
API runtime protection

AI SECURITY

OWASP Top 10 LLM attacks
Prompt injection
Data poisoning
Model theft

App Delivery

APPLICATION DELIVERY CONTROLLER (ADC)

CONTENT DELIVERY NETWORK (CDN)

API GATEWAY
API MANAGEMENT

API request routing
Dev portal
Rate limiting and accounting

AI / LLM PROXY

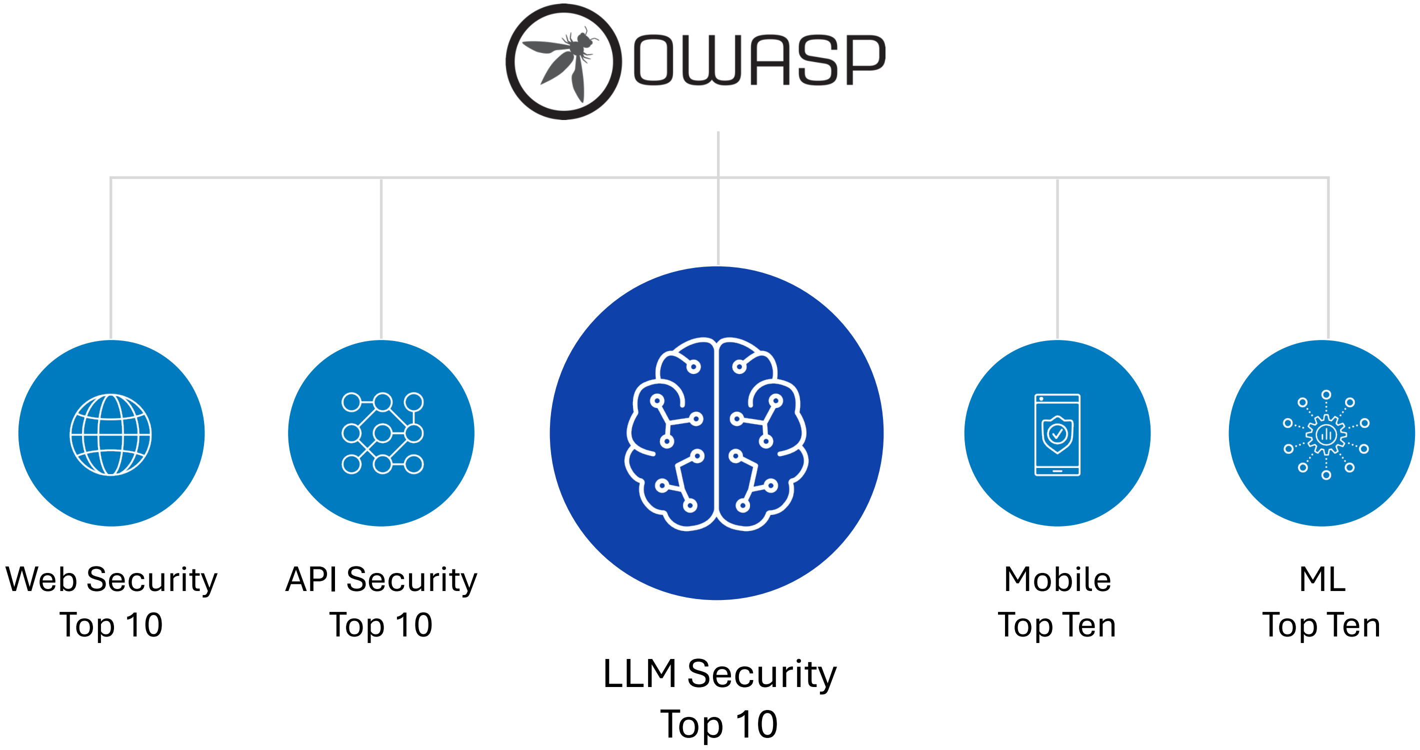
Deep LLM integrations
Prompt management
Rich observability
Token counting

Generative AI and API Security

API security at the foundation of AI security

> No matter how you’re using gen AI, at the end of the day, you’re calling an endpoint either with an SDK or a library or via a REST API.”

Mete Atamel
Developer Advocate, Google Cloud



OWASP

Web Security Top 10

API Security Top 10

LLM Security Top 10

Mobile Top Ten

ML Top Ten

F5 - Key Capabilities

Application and API Security

AI-powered protection for AI deployments with full API lifecycle protection

Traffic management and data ingestion

High-throughput connectivity across on-premises, cloud, and edge for AI inference and model training.

Operational simplicity and resource management

Simplified workload management and processing with a single platform that covers the entire application pipeline

Observability and Compliance

Precise monitoring and operational insights, coupled with stringent compliance measures

Top 5 Use Cases for the F5 Delivery & Security Network

1

Traffic Management for Data Ingest

High-performance load balancing to ingest data for multi-billion-parameter AI models

2

API Security

Automatically discover and secure all AI training and inference endpoints

3

Secure Multicloud Networking

High-throughput connectivity across on-prem, clouds, and edge for AI model training & inference

4

Distributed Inference

Run AI inference models anywhere cloud to edge with lifecycle management update, patch, and version manage centrally.

5

Secure AI App Delivery in Kubernetes

Simplify and streamline model operations and security at scale.

Key takeaways

Securing AI requires careful planning

New threat vectors

Protecting AI apps requires a security platform that can offer all existing app security tools (WAAP) plus additional protection against new attacks

Distributed architecture

Protecting an AI ecosystem requires a solution that allows to easily and safely interconnect all the AI system components across data center, cloud and edge

Evolving requirements

Look for solutions that allow you to evolve your architecture (inferencing, training, RAG, federated learning, etc.)

Get a free Web App Scanning