

EĞİTİM PROGRAMI

Agentic Siber Güvenlik Eğitimi

1 Günlük Program · Sıfırdan Başlayanlar İçin

AI ile Güvenlik · AI'nın Güvenliği

Agentic Siber Güvenlik Eğitimi

1 Günlük Program · Eğitim İçeriği

Eğitim Hakkında

Bu eğitim, yapay zekâ ajanlarının (AI agents) siber güvenlikte nasıl kullanıldığını ve bu ajanların kendilerinin nasıl bir saldırı yüzeyi oluşturduğunu, **sıfır teknik geçmişli olan** katılımcılara tek günde kazandırır.

Gün iki ana eksende ilerler:

- **Sabah — Temeller:** Yapay zekâ, LLM, agentic AI, MCP ve Claude Code kavramları sıfırdan, anlaşılır analogilerle anlatılır. Katılımcı gün sonunda bu kavramları başka birine anlatabilecek düzeye gelir.
- **Öğleden sonra — Uygulama:** Katılımcılar öğrendikleri her kavramı gerçek bir senaryo üzerinde uygular. Kod yazmadan, sadece Türkçe cümlelerle bir güvenlik ajanı kurar, çalıştırır, kırar ve korur.

Öğretim yaklaşımı: Her kavram **önce anlatılır, sonra uygulanır**. Katılımcı hiçbir şeyi ezberlemez; her fikri kendi eliyle deneyimleyerek öğrenir.

Kimler İçin?

- Siber güvenliğe ilgi duyan, ama teknik geçmişli olmayan profesyoneller
- Yapay zekânın güvenlik dünyasını nasıl değiştirdiğini merak eden yöneticiler ve ekipler
- Kariyerine güvenlik veya AI yönünde başlamak isteyen yeni başlayanlar

Ön Koşul

Programlama, Linux veya güvenlik bilgisi **gerekmez**. Tek gereken: bir dizüstü bilgisayar ve merak.

Katılımcı Ne Kazanır?

Gün sonunda katılımcı: 1. Yapay zekânın, LLM'in ve "agentic" olmanın ne demek olduğunu açıklayabilir. 2. Bir yapay zekâ ajanının nasıl "düşünüp" karar verdiğini anlatabilir. 3. MCP'nin ne işe yaradığını ve neden güvenlik açısından kritik olduğunu bilir. 4. Kod yazmadan, doğal dille bir güvenlik ajanı kurabilir. 5. Bir yapay zekâ ajanını kandırmanın (prompt injection) nasıl çalıştığını kendi eliyle görür. 6. Bir ajanı bu saldırılara karşı nasıl koruyacağını uygular. 7. Yapay zekâ destekli, kanıta dayalı bir güvenlik bulgusu yazabilir.

Günün Akışı

Zaman	Bölüm	İçerik
09:00 – 11:00	Temeller	Yapay zekâ, LLM, Claude, agentic AI, MCP, Claude Code
11:15 – 12:00	Uygulama 1	AI ile log analizi: İlk güvenlik ajanını kur
13:00 – 13:50	Uygulama 2	Yetkili keşif ve kapsam disiplini
14:00 – 14:50	Uygulama 3	Ajanı kırmak: Prompt injection
15:00 – 15:50	Uygulama 4	Ajanı korumak: Defence and Hardening
16:00 – 16:45	Bireysel Çalışma	Baştan sona tek başına bir olayı çöz
16:45 – 17:00	Kapanış	Değerlendirme, kariyer yolları, sonraki adımlar

BÖLÜM 1 — TEMELLER

Bu bölümün amacı, teknik olmayan bir katılımcının beş temel kavramı sağlam bir şekilde anlamasıdır. Her kavram günlük hayattan analogilerle anlatılır ve canlı örneklerle desteklenir.

1.1 Yapay Zekâ ve LLM Nedir?

Kavram

Yapay zekâ (AI), bilgisayarların normalde insan zekâsı gerektiren işleri yapabilmesidir: dil anlamak, örüntü tanımak, karar vermek. Bugün herkesin konuştuğu yapay zekânın kalbinde ise **LLM** vardır.

LLM (Large Language Model — Büyük Dil Modeli), devasa miktarda metinle eğitilmiş, **bir sonraki kelimeyi tahmin eden** bir sistemdir. Aslında yaptığı tek şey budur: verdiğiniz metnin ardından en olası kelimeyi, sonra bir sonrakini, sonra bir sonrakini üretmek. Ama bunu o kadar iyi yapar ki soru cevaplayabilir, kod yazabilir, metin özetleyebilir.

Analoji

LLM'i **her şeyi okumuş, çok zeki ama biraz fazla kendinden emin bir stajyer** gibi düşünün. İnternetteki kitapların, makalelerin, kodların büyük kısmını okumuştur. Ona bir soru sorduğunuzda, okuduğu her şeyden yola çıkarak en makul cevabı üretir. Ama emin olmadığı zaman bile durup "bilmiyorum" demez — güçlü bir tahmin yürütür ve size güvenle sunar.

Nasıl Çalışır? (3 Temel Kavram)

1. Bir sonraki kelime tahmini "İstanbul Türkiye'nin en kalabalık ___" cümlesini verirsiniz, model "şehridir" kelimesini yüksek olasılıkla tahmin eder. Tüm yetenekleri bu basit mekanizmanın üzerine kuruludur.

2. Context window (Bağlam penceresi) — Modelin masası Modelin "hafızası" o anki sohbetle sınırlıdır. Buna **context window** denir. Şöyle düşünün: modelin önünde bir **çalışma masası** var. Masanın üstünde ne varsa model onu görür ve kullanır. Masada olmayan hiçbir şey onun için mevcut değildir. Sohbet uzadıkça masa dolar; bir noktadan sonra eski bilgiler masadan düşer.

3. Halüsinasyon (Uydurma) Model bilmediği bir şeyi sorduğunuzda, çoğu zaman **kendinden emin bir şekilde uydurur**. Buna halüsinasyon denir. Bu bir hata değil, sistemin doğasıdır: model "en olası kelimeyi" üretir, "doğru cevabı" değil. Bu yüzden çıktısını her zaman doğrulamak gerekir.

Neden Önemli?

Bütün gün boyunca çalışacağımız ajanların beyni bir LLM'dir. Onun **ne yapabildiğini** (dil anlama, örüntü bulma, hızlı özet) ve **nerede tökezlediğini** (uydurma, kendinden emin yanlışlık) bilmek, bu araçları güvenle kullanmanın ilk şartıdır.

Bu bölümün altın kuralı: Model bir yanıt üretir; doğruluğunu garanti etmez. Doğrulamak sizin işiniz.

1.2 Claude Nedir?

Kavram

Claude, Anthropic firması tarafından geliştirilen bir yapay zekâ asistanıdır — yani belirli bir LLM ailesinin adıdır. Bu eğitimde uygulamalarımızı Claude üzerinde yapacağız.

Claude tek bir model değil, **farklı boyutlarda bir model ailesidir**. Büyük modeller daha güçlü ve daha derin düşünür; küçük modeller daha hızlı ve daha ekonomiktir. İşin ağırlığına göre uygun boyut seçilir — tıpkı bir işe kıdemli mi yoksa hızlı bir ekip mi göndereceğinize karar vermek gibi.

Claude'a Nasıl Erişilir?

Claude'a birden fazla yoldan ulaşılır; katılımcının bilmesi gereken üçü:

- **Sohbet arayüzü (claude.ai):** Web veya mobil üzerinden, ChatGPT benzeri bir sohbet penceresi. En basit kullanım.
- **API:** Yazılımcıların Claude'u kendi uygulamalarına gömmesini sağlayan teknik arayüz.
- **Claude Code:** Claude'un bilgisayarınızdaki dosyalara erişip iş yapabildiği, "ajan" gibi çalışan araç. Bu eğitimde uygulamaları bununla yapacağız (bkz. 1.5).

Neden Claude?

Claude, güvenlik ve güvenilirlik odağıyla tasarlanmıştır. Bir güvenlik eğitiminde, ajanın ne zaman durup insana danışması gerektiğini anlayan, kurallara uyan bir model kullanmak önemlidir. Uygulamalarda da göreceğimiz gibi, iyi bir ajanın en önemli özelliği "her istediğini yapması" değil, **doğru zamanda durup insana sormasıdır**.

1.3 Agentic AI Nedir? — Sohbet Botundan Ajana

Kavram

Buraya kadar anlattığımız LLM bir **sohbet botudur**: siz sorarsınız, o cevaplar, biter. **Agentic AI (Ajan yapay zekâ)** ise bunun bir üst seviyesidir: sadece cevaplamaz, **iş yapar**. Bir hedef verirsiniz, o hedefe ulaşmak için kendi kendine adımlar planlar, araçlar kullanır, sonuçları değerlendirir ve devam eder.

Sohbet Botu vs. Ajan

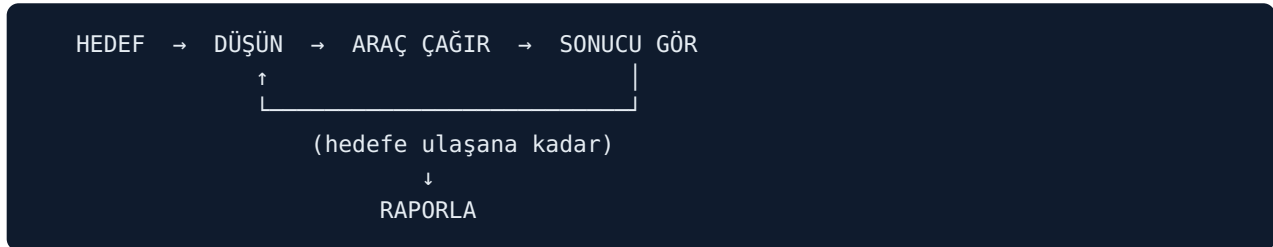
Sohbet Botu	Ajan
Tek soru sorarsınız	Bir hedef verirsiniz
Metinle cevap verir	Araçlar kullanarak iş yapar
Cevaplar, biter	Sonucu görür, bir sonraki adıma karar verir
Her seferinde sıfırdan başlar	Adımlar arası durumu hatırlar

Bir Ajanı Ajan Yapan 4 Şey

1. **Hedef (Goal):** Tek bir soru değil, ulaşılabilecek bir durum. Örnek: "Bu logları incele ve şüpheli olanları bul" — bu tek bir cevabı olan soru değil, bir görevdir.
2. **Araçlar (Tools):** Ajanın dünyayla etkileşim kurma yeteneği. Dosya okuyabilir, arama yapabilir, bir sisteme komut gönderebilir. Anlatmakla kalmaz, **yapar**.
3. **Döngü (Loop):** Ajanın çalışma ritmi. Bir adım atar, sonucu görür, buna göre bir sonraki adıma karar verir, tekrarlar.
4. **Hafıza (Memory):** Adımlar arasında durumu taşıma. İlk adımda bulduğunu üçüncü adımda kullanabilir.

Ajan Döngüsü — Kalbi Burası

Bir ajan şu döngüyü tekrar tekrar çevirir:



Örnek: "Bu logdaki şüpheli girişleri bul" dediğinizde ajan: 1. **Düşünür:** "Önce logu okumam lazım." 2. **Araç çağırır:** Dosya okuma aracını kullanır. 3. **Sonucu görür:** 300 satır log gelir. 4. **Düşünür:** "Başarısız girişleri saymam lazım." 5. **Araç çağırır:** Filtreleme yapar. 6. ...ve hedefe ulaşana kadar döngüyü çevirir, sonunda raporlar.

Kritik Ayrım — Güvenliğin Temeli Burası

Ajan bir aracı **kendisi çalıştırmaz**. Aracın çalıştırılmasını **talep eder**; bir program bu talebi görür, onaylar veya reddeder. Yani ajanla gerçek dünya arasında her zaman bir onay katmanı vardır.

Bu cümle çok önemli, çünkü **ajan güvenliğinin tamamı** bu ayrımın üzerine kuruludur. Ajanın "yapmak istediği" ile "gerçekten yapabildiği" arasındaki farkı kontrol edebildiğimiz için ajanları güvenle kullanabiliriz. Öğleden sonra bu katmanın ne kadar kritik olduğunu kendi elinizle göreceksiniz.

Neden Güvenlik İçin Önemli? — İki Yön

Ajanlar güvenlik dünyasına iki yönden girer: - **AI ile güvenlik:** Ajanları savunma ve keşif işlerinde bir güç çarpanı olarak kullanmak. (Uygulama 1 ve 2) - **AI'ın güvenliği:** Ajanların kendisinin yeni bir saldırı yüzeyi olması. (Uygulama 3 ve 4)

Bu eğitim **her iki yönü de** kapsar.

1.4 Araçlar ve MCP Nedir?

Kavram

Bir ajanın "iş yapabilmesi" için **araçlara (tools)** ihtiyacı vardır. Araç, ajanın dünyayla etkileşim kurma yoludur: bir dosyayı okumak, bir veritabanını sorgulamak, bir e-posta göndermek. Ajanın bir araca "bunu çalıştır" demesine **function calling (fonksiyon çağırma)** denir.

Peki bu araçlar ajana nasıl bağlanır? İşte burada **MCP** devreye girer.

MCP Nedir?

MCP (Model Context Protocol — Model Bağlam Protokolü), yapay zekâ ajanlarını dış araçlara ve veri kaynaklarına bağlamak için geliştirilmiş **ortak bir standarttır**.

Analoji — USB-C Girişi

MCP'yi telefon ve bilgisayarlardaki **USB-C girişi** gibi düşünün. Eskiden her cihazın kendine özel bir şarj kablosu vardı; her biri için ayrı kablo taşımak gerekirdi. USB-C tek bir standart getirdi: artık tek tip girişle her şeyi bağlıyorsunuz.

MCP de yapay zekâ için aynısını yapar. Her araç için ajana özel "kablo" çekmek yerine, araçlar **tek tip bir prize** takılır. Ajan bu prize takılı araçları **keşfeder** ve kullanır.

MCP Ne Çözer?

- **Standartlaşma:** Her araç aynı dille konuşur, her şeyi baştan yazmak gerekmez.
- **Keşfedilebilirlik:** Ajan, prize hangi araçların takılı olduğunu kendisi görebilir.
- **Yeniden kullanım:** Bir kez yazılan bir MCP aracı, farklı ajanlarla tekrar tekrar kullanılır.

MCP Ne ÇÖZMEZ? — Güvenlik Uyarısı

MCP araçları bağlamayı kolaylaştırır, ama **taktığınız aracın niyetinin iyi olduğunu garanti etmez**. Prize taktığınız şey zehirli olabilir. Kötü niyetli bir araç, açıklamasının içine gizli talimatlar koyarak ajanı kandırabilir.

Öğleden sonra Uygulama 3'te, ajanın okuduğu bir içeriğin nasıl bir silaha dönüştüğünü kendi elinizle göreceksiniz. Şimdilik şu cümleyi aklınızda tutun: **Prize taktığınız şeyin içeriği güvenilir değildir.**

1.5 Claude Code Nedir ve Kod Bilmeyen Biri Nasıl Kullanır?

Kavram

Claude Code, Claude'un bilgisayarınızın içinde iş yapabilen "ajan" halidir. Dosyalarınızı okur, düzenler, komut çalıştırır ve MCP araçlarıyla dış sistemlere bağlanır. Kısacası, 1.3'te anlattığımız "ajan döngüsünü" gerçek bir bilgisayarda çalıştıran araçtır.

"Ama Ben Kod Bilmiyorum!" — Sorun Değil

Bu eğitimin en önemli mesajı burada: **Claude Code'u kullanmak için kod bilmeniz gerekmez.**

Claude Code ile **düz cümlelerle** konuşursunuz. Ona ne istediğinizi Türkçe (veya İngilizce) anlatırsınız, o gerisini halleder. Teknik komutları, dosya işlemlerini, tüm karmaşık kısmı **o** **yapar.**

Örnek — katılımcının yazacağı gerçek bir komut:

"Bu klasördeki dosyaları listele ve her birinin ne olduğunu bir cümleyle açıkla."

Katılımcı bir tek teknik komut yazmaz. Sadece isteğini söyler. Claude Code: 1. Klasörü okur (araç çağırır), 2. Dosyaları listeler (sonucu görür), 3. Her birini açıklar (raporlar).

İşte katılımcı, ilk beş dakikada ajan döngüsünü **çalışırken canlı görür.**

Bu Eğitimde Kural

Bugün tek bir teknik komut yazmayacaksınız. Sadece ne istediğinizi cümlelerle anlatacaksınız.

Bu kural, sıfırdan başlayan katılımcının önündeki en büyük engeli (teknik korku) daha ilk dakikada kaldırır. Katılımcı "kod yazamıyorum" diye takılmaz; çünkü kod yazması gerekmiyor.

İlk Temas — Nasıl Başlıyoruz?

Uygulamalara geçmeden önce herkes birlikte, güvenli bir "merhaba dünya" adımı atar: yukarıdaki basit komutu kendi ekranında çalıştırır ve ajanın kendi kendine araç çağırdığı anı görür. Eğitmen bu anı isimlendirir: *"İşte döngü bu. Gördünüz mü? Plan yaptı, araç kullandı, sonucu gördü, size anlattı."*

Bu andan sonra katılımcı, ajanın ne olduğunu **teoride değil, ekranında** bilir.

1.6 Büyük Resim — İki Yön

Temeller bölümünü kapatırken, günün geri kalanının haritasını çizelim:

	AI ile Güvenlik	AI'nın Güvenliği
Fikir	Ajanı bir güvenlik aracı olarak kullan	Ajanın kendisi bir zafiyet
Örnek	Logları ajanla tara, olayları hızlıca bul	Ajanı kandırıp gizli veriyi sızdır
Uygulama	Uygulama 1 & 2	Uygulama 3 & 4

Ve gün boyunca 5 kez tekrarlanacak üç kural:

1. **AI beyindir. İnsan onaylayıcıdır.** — Ajan önerir, insan karar verir. Hiçbir şey insan onayı olmadan gerçek bir sisteme ulaşmaz.
2. **Kanıtı olmayan bulgu, bulgu değildir.** — Ajan "kritik" dedi diye kritik olmaz; kanıt gözünüzle görene kadar hiçbir şey doğru değildir.
3. **Bu problemin tam çözümü yok. Alan hâlâ açık.** — Bazı saldırıların bugün bile kesin çözümü yoktur. İşte bu yüzden bu bir meslektir.

KAPANIŞ

Değerlendirme ve Geri Bildirim

Katılımcılar, iki örnek raporun canlı olarak değerlendirildiği kısa bir oturumla günü tamamlar. Bu bölüm, iyi bir güvenlik raporunun neye benzediğini somut olarak gösterir.

Dürüst Tablo — Yapay Zekâ Bugün Nerede?

Eğitim, gerçekçi bir bakışla kapanır:

Gerçekten İyi Yaptığı	Sadece Pazarlama	Hâlâ İnsan İşi
Log tarama, hızlı triyaj	"Tam otomatik güvenlik"	Raporun sorumluluğu
Örüntü bulma, korelasyon	"İnsana gerek kalmadı"	Kapsam ve etik kararlar
İlk taslağı yazma	"Tek tıkla sızma testi"	Nihai onay ve yargı

Kariyer Yolları

- **SOC Analisti → Tespit Mühendisi:** Log analizi ve kural yazımı.
- **Sızma Testçisi → AI Red Team:** Sistemleri kırmayı öğrendiniz, şimdi yapay zekâ sistemlerini test edin.
- **Yazılımcı → AI Güvenlik Mühendisi:** Güvenli ajan tasarımı.

30 Günlük Yol Haritası

Katılımcılara, öğrendiklerini pekiştirmeleri için haftalık bir plan verilir: uygulamaları yardımsız tekrar etmek, OWASP LLM Top 10 listesini incelemek, açık prompt injection alıştırmaamlarını tamamlamak ve bir ajanı baştan sona savunmak.

Eğitim Sonunda Katılımcı...

- Yapay zekâ, LLM, agentic AI, MCP ve Claude Code kavramlarını başkasına anlatabilir.
- Kod yazmadan, doğal dille bir güvenlik ajanı kurabilir.
- Bir ajanı hem kullanabilir, hem kırabilir, hem koruyabilir.
- Kanıta dayalı, profesyonel bir güvenlik bulgusu yazabilir.
- Ve en önemlisi: yapay zekânın nerede güçlü, nerede tehlikeli olduğunu — teoride değil, kendi eliyle — bilir.

AI beyindir. İnsan onaylayıcıdır.