Acil durdurma düğmeleri, denetçiler ve egemen bir ittifak

CrowdStrike yapay zeka ajanları için bir acil durdurma düğmesi sundu, ABD'li OCC banka denetçilerine yapay zekayı hatanın bedeline göre değerlendirmelerini söyledi, Bull ve Schwarz Digits egemen yapay zeka için güçlerini birleştirdi, Capgemini de finansta çok ajanlı yapay zekanın ne kadar yaygınlaştığını ölçtü.

Son haftaların ana teması, yani ajanları yalnızca devreye almak değil denetlemek, laboratuvarlardan ürünlere ve denetim otoritelerine taşındı. Büyük bir güvenlik şirketi artık tüm ajanları tek hamlede durdurmayı sağlayan bir çözüm satıyor; ABD'nin başlıca banka denetim kurumu da denetçilerine otonom ödeme ajanlarını en yüksek risk grubuna koyan bir çerçeve verdi. Avrupa'da bir Fransız ve bir Alman şirketi egemen yapay zeka altyapısını birlikte kurmak için anlaştı; yeni bir araştırma ise bankaların ve sigortacıların çoğunun çok ajanlı sistemleri denediğini, ancak çok azının bunları ölçekli çalıştırdığını gösteriyor.

AJANLAR VE ARAÇLAR

CrowdStrike, yapay zeka ajanları için acil durdurma düğmesi sundu

CrowdStrike, 3 Kasım'daki Fal.Con Europe konferansında güvenlik ekiplerinin ajanların yapay zeka modellerine erişimini ağ geçidinde tek hamlede kesmesini sağlayan bir acil durdurma düğmesi tanıttı; ajanları uç cihazlarda durdurma özelliğinin ise daha sonra geleceği belirtiliyor. Şirket ayrıca her ajanın olağan davranışını öğrenip herhangi bir kural ihlali olmasa bile sapmaları işaretleyen bir tespit özelliği duyurdu; ancak bu özellik henüz erken aşamada.

Neden önemli: Denetim otoriteleri, kurumların bir yapay zeka sistemini hızla durdurabildiğini göstermesini giderek daha fazla bekliyor. Canlı sistemlere dokunan her ajan için test edilmiş bir kapatma yolu, yedekleme kadar temel bir gereklilik haline geliyor.

Bizim görüşümüz: Merkezi bir durdurma düğmesi işe yarar; ama en iyi sonucu güvenli biçimde durmak üzere tasarlanmış ajanlarla birlikte verir. Ajan iş akışlarını, görevin ortasında durdurulduğunda kayıtların tutarlı kalacağı ve yarım kalan adımın açıkça kayda geçeceği şekilde kuruyoruz.

Kaynak: CrowdStrike

Adobe, kurumsal ajanını ChatGPT, Claude, Gemini ve Copilot'un içine taşıdı

Adobe 4 Kasım'da, pazarlama ve müşteri deneyimi işleri için geliştirdiği CX Enterprise Coworker ajanının artık ChatGPT, Claude, iş kullanımına yönelik Gemini, Microsoft Copilot ve Databricks Genie içinde çalıştığını açıkladı; Amazon, IBM ve Snowflake platformları da sırada. Her asistanda, çalışanların kullanabilmesi için uygulamanın önce bir yönetici tarafından onaylanması gerekiyor.

Neden önemli: Kurumsal yazılımlar çalışanlara kendi ekranları yerine genel amaçlı asistanlar üzerinden ulaşmaya başlıyor. Bu da erişim kontrolünün ve denetim kaydının bir kısmını asistan platformuna taşıyor.

Bizim görüşümüz: Bir tedarikçi ajanını asistanın içinde onaylamadan önce hangi verilere erişebildiğini, işlemlerinin nerede kaydedildiğini ve bu kayıtların kendi izleme sistemlerinize ulaşıp ulaşmadığını kontrol edin. Bunu sormanın doğru anı yönetici onayı aşaması.

Kaynak: Adobe

KURUMSAL KULLANIM

Capgemini: Bankaların ve sigortacıların çoğu çok ajanlı yapay zekayı deniyor, azı ölçekli kullanıyor

Capgemini'nin 5 Kasım'da yayımladığı araştırmaya göre ankete katılan banka ve sigorta şirketlerinin %59'u çok ajanlı yapay zeka sistemlerini pilot olarak deniyor ya da devreye alıyor. Ajanların birimler arasında birlikte çalıştığı kurum çapında orkestrasyona ulaşanların oranı ise yalnızca %10 civarında. Araştırma, bu dönüşümün ana itici gücü olarak ölçülebilir iş getirisini gösteriyor.

Neden önemli: Finansal hizmetlerde yapay zekanın asıl zorluğu artık pilot ile üretim arasındaki açık. Tek bir ajandan çok sayıda ajana geçmek; sahiplik, görev devri ve denetim konusunda yeni sorular doğuruyor.

Bizim görüşümüz: Orkestrasyon en sık ajanlar arasındaki görev devirlerinde aksar. Ajanları tek bir iş akışında birleştirmeden önce her ajanın neyi devrettiğini, sonucun kime ait olduğunu ve insan onayının nerede devreye girdiğini tanımlıyoruz.

Kaynak: Capgemini

REGÜLASYON VE YÖNETİŞİM

ABD'li OCC, banka denetçilerine riski hatanın bedeline göre ölçen bir yapay zeka çerçevesi verdi

ABD Para Denetim Ofisi (OCC), 4 Kasım'da bankaların yapay zeka kullanımını inceleyen denetçiler için ilk denetim çerçevesini yayımladı. Çerçeve, denetçilerden bir yapay zeka kullanımını hatanın neye mal olacağına, ne kadar geri alınabilir olduğuna ve ne ölçüde insan gözetimi içerdiğine göre değerlendirmelerini istiyor; otonom ödeme ajanlarını ve anlık kredi kararlarını en yüksek risk grubuna yerleştiriyor. OCC, çerçevenin denetçiler için bir kaynak olduğunu ve bankalara yeni yükümlülük getirmediğini belirtiyor.

Neden önemli: Yeni kurallar olmasa da bu çerçeve denetçilerin neye bakacağını gösteriyor: Devredilen yetkiler, işlem limitleri, izleme ve bir ajanı kapatmanın test edilmiş bir yolu. Avrupalı denetim otoriteleri de benzer soruları soruyor.

Bizim görüşümüz: Sonuca göre risk sınıflandırması, nereden başlanacağına karar vermenin pratik bir yolu. Her ajanın işlemlerini bir hatanın neye mal olacağına göre eşleştiriyor, en sıkı kontrolleri ödemeler gibi geri alınamayan işlemlere uyguluyoruz.

Kaynak: American Banker

EGEMEN YAPAY ZEKA VE ALTYAPI

Bull ve Schwarz Digits, egemen yapay zeka altyapısını birlikte kurmak için anlaştı

Fransız Bull ve Lidl'in bağlı olduğu grubun bilişim kolu olan Alman Schwarz Digits, 3 Kasım'da regüle ve stratejik sektörler için Avrupa yargı yetkisi altında çalışacak yapay zeka ve bulut hizmetleri geliştirmek üzere bir mutabakat zaptı imzaladı. Schwarz Digits STACKIT bulutunu, Bull ise yüksek başarımlı hesaplama uzmanlığını ve Angers'deki fabrikasında üretilen sunucuları sağlayacak. Anlaşma ön niteliğinde; mali koşullar ya da teslim tarihleri açıklanmadı.

Neden önemli: Avrupa'daki regüle kurumlar için yerel sahiplik, yerel donanım ve yapay zeka kapasitesini bir arada sunan sağlayıcı sayısı az. Fransız-Alman ortaklığıyla kurulacak bir altyapı, ürün haline geldiğinde seçenekleri genişletecek.

Bizim görüşümüz: Mutabakat zaptı bir hizmet değil, bir yön göstergesidir. Egemen seçenekleri değerlendirirken bugün gerçekten sunulan ve sözleşmeyle taahhüt edilen hizmetlere bakıyor, sistemleri ileride daha iyi bir Avrupa seçeneğine geçilebilecek şekilde tasarlıyoruz.

Kaynak: The Next Web

BÖLGESEL GÜNDEM

CrowdStrike, Suudi Arabistan'da ülke içi bulut hizmetini açtı

CrowdStrike 4 Kasım'da, Falcon güvenlik platformunun artık Suudi Arabistan'da ülke içinde çalıştığını; müşteri telemetrisi, iş yükleri ve yönetim altyapısının yerel olarak barındırıldığını açıkladı. Şirkete göre yapı, Suudi siber güvenlik ve veri koruma kurallarına uygun şekilde tasarlandı ve hemen kullanılabilir durumda. Bu, CrowdStrike'ın ocak ayında Suudi Arabistan, Hindistan ve BAE için duyurduğu ülke içi bulutların ilki.

Neden önemli: Güvenlik araçları çok sayıda hassas veriyi görür. Suudi Arabistan'daki bankalar ve enerji şirketleri için bu telemetrinin ülke içinde kalması, veri yerleşimine ilişkin yaygın bir itirazı ortadan kaldırıyor.

Bizim görüşümüz: Veri yerleşimi yalnızca model için değil, yapay zekanın etrafındaki araçlar için de geçerli. Körfez'de ülke içi yapay zeka kurulumları tasarlarken kayıt, izleme ve güvenlik telemetrisini de veri yerleşimi planına dahil ediyoruz.

Kaynak: CrowdStrike

TAKİP ETTİKLERİMİZ

Microsoft, Suudi Arabistan Doğu bulut bölgesinin bu ay müşterilere açılacağını duyurmuştu. İngiltere Bilgi Komiserliği'nin (ICO) ajan tabanlı yapay zekaya ilişkin görüş toplama süreci 20 Kasım'da sona eriyor; Katar'ın Future AI Summit etkinliği ise 23 Kasım'da Doha'da yapılacak. Avrupa'da Google 17 Kasım'da Brüksel'de bir egemen yapay zeka zirvesi, OVHcloud da 19 Kasım'da Paris'te yapay zeka zirvesini düzenliyor.

Bunu kurumunuzda uygulamak ister misiniz?

Bankacılık, sigorta, üretim ve enerji için üretim ortamına hazır yapay zeka ajanları geliştiriyoruz. Çalışan bir POC genellikle bir ila iki hafta sürer.

agenticdynamic.com'da okuyun **Bizimle görüşün →**