T.C.
IŞIK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI
SİBER GÜVENLİK PROGRAMI

Nezih Mahmut ÜNAL

KAMUYA AÇIK BÜYÜK DİL MODELLERİ İLE BAĞLAM DUYARLI SİBER RİSK DEĞERLENDİRMESİ: UZMAN DOĞRULAMALI BİR ÇERÇEVE VE İNSAN–YAPAY ZEKÂ KARŞILAŞTIRMASI

DANIŞMAN
Dr. Öğr. Üyesi Barış ÇELİKTAŞ

İSTANBUL, Ocak 2026

T.C.
IŞIK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
SİBER GÜVENLİK PROGRAMI

Nezih Mahmut ÜNAL

(23SIBE5011)

KAMUYA AÇIK BÜYÜK DİL MODELLERİ İLE BAĞLAM DUYARLI SİBER RİSK DEĞERLENDİRMESİ: UZMAN DOĞRULAMALI BİR ÇERÇEVE VE İNSAN–YAPAY ZEKÂ KARŞILAŞTIRMASI

DANIŞMAN

Dr. Öğr. Üyesi Barış ÇELİKTAŞ

İSTANBUL, Ocak 2026

T.C.
IŞIK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI
SİBER GÜVENLİK PROGRAMI

Nezih Mahmut ÜNAL

(23SIBE5011)

KAMUYA AÇIK BÜYÜK DİL MODELLERİ İLE BAĞLAM DUYARLI SİBER RİSK DEĞERLENDİRMESİ: UZMAN DOĞRULAMALI BİR ÇERÇEVE VE İNSAN–YAPAY ZEKÂ KARŞILAŞTIRMASI

Tezin Savunulduğu Tarih: 19.01.2026

Tez Danışmanı: Dr. Öğr. Üyesi Barış ÇELİKTAŞ / Işık Üniversitesi

Diğer Jüri Üyeleri: Dr. Öğr. Üyesi Emine EKİN / Işık Üniversitesi

Dr. Öğr. Üyesi Mehmet Tahir SANDIKKAYA / İTÜ

İSTANBUL, Ocak 2026

ÖZET

KAMUYA AÇIK BÜYÜK DİL MODELLERİ İLE BAĞLAM DUYARLI SİBER RİSK DEĞERLENDİRMESİ: UZMAN DOĞRULAMALI BİR ÇERÇEVE VE İNSAN–YAPAY ZEKÂ KARŞILAŞTIRMASI

Geleneksel siber risk değerlendirme metodolojileri kritik bir ikilemle karşı karşıyadır: Bu yöntemler ya nicel ancak statik ve bağlamdan bağımsızdır (Örn: CVSS) ya da bağlama duyarlı ancak yoğun emek gerektiren ve özneldir (Örn: NIST SP 800-30). Sonuç olarak kuruluşlar, risk değerlendirme süreçlerini gelişen tehditlerin hızına uyum sağlayacak şekilde ölçeklendirmekte zorlanmaktadır. Bu çalışma; uzman bilgisini işlevsel hale getirmek amacıyla kamuya açık Büyük Dil Modellerinin (LLM) akıl yürütme yeteneklerinden yararlanan, otomatik ve bağlama duyarlı bir risk değerlendirme çerçevesi sunmaktadır. Karmaşık "kapalı kutu" (black-box) makine öğrenmesi modellerinin aksine, önerilen yaklaşım yapay zekanın akıl yürütme sürecini şeffaf bir Dinamik Metrik Motoruna dayandırmaktadır. Bu motorun ağırlıkları, 101 siber güvenlik profesyoneli ile gerçekleştirilen bir anket çalışmasından Sıralı Derece Ağırlık Merkezi (Rank Order Centroid - ROC) yöntemi kullanılarak elde edilmiştir. Geliştirilen çerçeve, 15 farklı gerçek dünya zafiyet senaryosu (C_1--C_{15}) ve üç ek duyarlılık stres testi (C_{16}--C_{18}) içeren karşılaştırmalı bir çalışma aracılığıyla değerlendirilmiştir. Doğrulama senaryoları, on kıdemli uzmandan oluşan bir grup ve iki modern LLM ajanı (GPT-4o ve Gemini 2.0 Flash) tarafından bağımsız olarak analiz edilmiştir. Elde edilen sonuçlar, LLM tabanlı ajanların oldukça güvenilir bir uzman temel çizgisine (Cronbach's α = 0,996) karşı, insan medyanıyla yakından uyumlu bir puanlama tutarlılığı (Pearson r değeri 0,9390 ile 0,9717 ; Spearman ρ değeri 0,8472 ile 0,9276 aralığında) sergilediğini göstermiştir. Ayrıca sistem,

değerlendirme döngü süresini 100 kattan fazla azaltmıştır (vaka başına ortalama 6 dakikalık insan süresine karşı 4 saniyenin altı). Dahası, özel bir bağlam duyarlılık analizi (C_{13}--C_{15}); çerçevenin, özdeş teknik zafiyetler için risk skorlarını kurumsal bağlama (örneğin KOBİ'ye karşı Kritik Altyapı) göre uyarlayabildiğini kanıtlamıştır. Genel olarak bu bulgular, ticari olarak erişilebilen LLM'lerin uzmanlarca doğrulanmış metrik şemalarıyla sınırlandırıldığında; tekrarlanabilir, doğru ve gerçek zamanlı risk değerlendirmelerini destekleyebileceğini ortaya koymaktadır.

Anahtar Kelimeler: Siber Risk Değerlendirmesi, Büyük Dil Modelleri (LLM), Sıralı Derece Ağırlık Merkezi (ROC), Otomatik Risk Puanlama, İnsan-YZ Karşılaştırması.

ABSTRACT

LLM-ASSISTED CONTEXT-AWARE CYBER RISK ASSESSMENT: AN EXPERT-CALIBRATED FRAMEWORK AND SCORING

Traditional cyber risk assessment methodologies face a critical dilemma: they are either quantitative yet static and context-agnostic (e.g., CVSS), or context-aware yet highly labor-intensive and subjective (e.g., NIST SP 800-30). Consequently, organizations struggle to scale risk assessment to match the pace of evolving threats. This paper presents an automated, context-aware risk assessment framework that leverages the reasoning capabilities of publicly available Large Language Models (LLMs) to operationalize expert knowledge. Unlike complex black-box machine learning models, our approach anchors the AI's reasoning to a transparent Dynamic Metric Engine, with weights derived using the Rank Order Centroid (ROC) method from a survey of 101 cybersecurity professionals. We evaluated the framework through a comparative study involving 15 diverse real-world vulnerability scenarios (C_1--C_{15}) and three supplementary sensitivity stress tests (C_{16}--C_{18}). The validation scenarios were independently assessed by a cohort of ten senior human experts and two state-of-the-art LLM agents (GPT-4o and Gemini 2.0 Flash). The results show that the LLM-driven agents achieve scoring consistency closely aligned with the human median (Pearson r ranging from 0.9390 to 0.9717, Spearman ρ from 0.8472 to 0.9276) against a highly reliable expert baseline (Cronbach's α=0.996\$), while reducing the assessment cycle time by more than 100\$\\times\$ (averaging under 4 seconds per case vs. a human average of 6 minutes). Furthermore, a dedicated context sensitivity analysis (C_{13}--C_{15}) indicates that the framework adapts risk scores based on organizational context (e.g., SME vs. Critical Infrastructure) for identical technical vulnerabilities. Overall, these

findings suggest that commercially available LLMs, when constrained by expert-validated metric schemas, can support reproducible, accurate, and real-time risk assessments.

Keywords: Cyber risk assessment, Large Language Models (LLMs), Generative AI, Automated Risk Scoring, Human-AI Comparison.

TEŞEKKÜR

Bu tezin planlanması, yürütülmesi ve tamamlanması sürecinde değerli vaktini ayıran, vizyonu ve akademik rehberliği ile bana her zaman yol gösteren değerli danışmanım Dr. Öğr. Üyesi Barış ÇELİKTAŞ'a en içten şükranlarımı sunarım.

Eğitim hayatım boyunca akademik gelişimime katkı sağlayan Işık Üniversitesi siber güvenlik programındaki tüm değerli hocalarıma teşekkür ederim.

Son olarak, bu uzun ve yoğun çalışma sürecinde sabırları, fedakarlıkları ve sonsuz destekleri ile her zaman yanımda olan sevgili eşim Emel'e ve oğlum Kaan Alp'e gönülden teşekkür ederim.

Nezih Mahmut ÜNAL

İÇİNDEKİLER

ŞEKİLLER LİSTESİ

TABLOLAR LİSTESİ

KISALTMALAR LİSTESİ

AHP-TOPSIS : Analytic Hierarchy Process - Technique for Order of Preference by Similarity to Ideal Solution (Analitik Hiyerarşi Süreci - İdeal Çözüme Benzerliğe Göre Tercih Sıralama Tekniği)

ANN : Artificial Neural Networks (Yapay Sinir Ağları)

API : Application Programming Interface (Uygulama Programlama Arayüzü)

ARA : Adversarial Risk Analysis (Hasmane Risk Analizi)

BAĞ : Varlık/Hedef Bağlam Farkındalığı

BT/OT : Bilgi Teknolojileri / Operasyonel Teknolojiler (IT/OT - Information Technology / Operational Technology)

CISO : Chief Information Security Officer (Bilgi Güvenliği Üst Düzey Yöneticisi)

CTI : Cyber Threat Intelligence (Siber Tehdit İstihbaratı)

CVSS : Common Vulnerability Scoring System (Ortak Zafiyet Puanlama Sistemi)

CWSS : Common Weakness Scoring System (Ortak Zayıflık Puanlama Sistemi)

ETK : Teknik Etki

FAIR : Factor Analysis of Information Risk (Bilgi Riskinin Faktör Analizi)

FIRST : Forum of Incident Response and Security Teams

HITL : Human-in-the-Loop (İnsan Denetimli Süreç / Döngüdeki İnsan)

IDO : Intent-Driven Operations (Niyet Odaklı Operasyonlar)

IoHT : Internet of Health Things (Sağlık Nesnelerinin İnterneti)

IoT : Internet of Things (Nesnelerin İnterneti)

İŞ : İş/Mevzuat Etkisi

JSON : JavaScript Object Notation (JavaScript Nesne Notasyonu)

KOBI : Küçük ve Orta Büyüklükteki İşletmeler

KUR : Kurumsal Faktörler

LIME : Local Interpretable Model-agnostic Explanations (Yerel Yorumlanabilir Modelden Bağımsız Açıklamalar)

LLM : Large Language Models (Büyük Dil Modelleri)

MEA : Mean Absolute Error (Ortalama Mutlak Hata)

MFA : Multi-Factor Authentication (Çok Faktörlü Kimlik Doğrulama)

ML : Machine Learning (Makine Öğrenmesi)

MÖ : Makine Öğrenmesi

NLP : Natural Language Processing (Doğal Dil İşleme)

OCTAVE : Operationally Critical Threat, Asset, and Vulnerability Evaluation (Operasyonel Olarak Kritik Tehdit, Varlık ve Zayıflık Değerlendirmesi)

OWASP : Open Web Application Security Project

ÖMA : Özelleştirilebilir Metrik Ağırlıklandırma

PCI DSS : Payment Card Industry Data Security Standard (Ödeme Kartı Endüstrisi Veri Güvenliği Standardı)

PII : Personally Identifiable Information (Kişisel Olarak Tanımlanabilir Bilgi)

RAG : Retrieval-Augmented Generation (Veri Erişimi ile Artırılmış Üretim)

RCE : Remote Code Execution (Uzaktan Kod Çalıştırma)

ROC : Rank Order Centroid (Sıralı Ağırlık Merkezi)

SAL : Saldırgan Modelleme

SHAP : Shapley Additive Explanations (Shapley Katkılı Açıklamalar)

SLM : Small Language Model (Küçük Dil Modeli)

SME : Small and Medium-sized Enterprises (KOBİ - Küçük ve Orta Büyüklükteki İşletmeler)

SOC : Security Operations Center (Siber Güvenlik Operasyon Merkezi)

SÖM : Sömürülebilirlik

TTP : Tactics, Techniques, and Procedures (Taktikler, Teknikler ve Prosedürler)

VPN : Virtual Private Network (Sanal Özel Ağ)

VPS : Virtual Private Server (Sanal Özel Sunucu)

WAF : Web Application Firewall (Web Uygulaması Güvenlik Duvarı)

WF : Weight Factor (Ağırlık Faktörü)

XAI : Explainable Artificial Intelligence (Açıklanabilir Yapay Zeka)

YZ : Yapay Zeka

BÖLÜM 1

1. GİRİŞ

Artan dijital bağımlılık ve tehdit unsurların artan sofistike yöntemleri, siber güvenlik risk yönetimini kurumsal dayanıklılığın temel taşı haline gelmiştir. Bu süreçte, veriye dayalı ve uyarlanabilir yaklaşımlar etkili güvenlik stratejilerinin temel sağlayıcıları olarak kabul edilmektedir (Yu vd., 2024). Ancak mevcut birçok risk değerlendirme metodolojisi, zafiyetlerin karmaşık bağımlılık yapılarını yakalamakta zorlanmakta; bu durum kapsamlı ve güvenilir risk değerlendirme kabiliyetlerini kısıtlamaktadır (He vd., 2019). Söz konusu sınırlamalar, dinamik kaynak tahsisi ve çoklu kiracılık (multi-tenancy) gibi unsurların risk değerlendirme sürecine ek belirsizlik ve karmaşıklık getirdiği bulut bilişim ortamlarında daha da belirginleşmektedir (Drissi vd., 2025).

Bu zorlukları aşmak amacıyla çeşitli metodolojiler önerilmiştir. Yaygın olarak benimsenen Ortak Zafiyet Puanlama Sistemi (CVSS) (Forum of Incident Response and Security Teams [FIRST], 2023), Bilgi Riski Faktör Analizi (FAIR) (Freund ve Jones, 2014) ve OWASP Risk Derecelendirmesi gibi puanlama yaklaşımları, zafiyetlerin nicelleştirilmesi için standart mekanizmalar sunmaktadır. Eş zamanlı olarak; ISO/IEC 27005 (International Organization for Standardization [ISO], 2022), ISO 31000 (International Organization for Standardization [ISO], 2018), NIST SP 800-30 (National Institute of Standards and Technology [NIST], 2012) ve OCTAVE (Alberts ve Dorofee, 2002) gibi kapsamlı risk değerlendirme çerçeveleri ve standartlar yapılandırılmış yönetişim süreçleri sağlamaktadır. Bu yaklaşımlar alanı önemli ölçüde geliştirmiş olsa da, kritik bir "ölçeklenebilirlik-hassasiyet" ikilemi ile karşı karşıyadırlar (Yu vd., 2024; He vd., 2019). CVSS gibi statik puanlama sistemleri hızlıdır ancak bağlamdan bağımsızdır. Kurumsal savunmaları göz ardı ettikleri için düşük öncelikli sorunları sıklıkla yüksek riskli olarak işaretlemektedirler (Yu vd., 2024; Drissi vd., 2025). Aksine, nitel çerçeveler derin bir bağlamsal farkındalık

sağlasa da büyük ölçüde manuel uzman görüşüne dayanmaktadır. Bu bağımlılık, otomatik tehditler karşısında öznelliğe, tutarsızlığa ve yavaş geri dönüş sürelerine yol açabilmektedir. Sonuç olarak kuruluşlar, risk değerlendirme çıktıları ile iyileştirme önceliklendirmesi için gereken gerçek zamanlı kararlar arasında bir kopukluk yaşamaktadır (Yu vd., 2024).

Büyük Dil Modellerinin (LLM) ortaya çıkışı bu konuda yeni bir yol sunmaktadır. Geleneksel statik algoritmaların aksine LLM'ler; yapılandırılmamış siber tehdit istihbaratını (CTI) yorumlayabilmekte, kurumsal bağlamı sürece dahil edebilmekte ve bir insan analistine benzer şekilde risk seviyelerini tahmin edebilmektedir. Ancak, yalnızca "kapalı kutu" yapay zeka modellerine güvenmek; halüsinasyon, sınırlı açıklanabilirlik ve tutarsızlık gibi yeni riskleri beraberinde getirmektedir. Bu opaklığı gidermek amacıyla "Sistemik Şeffaflık" kavramı sunulmaktadır. Burada önerilen çerçevede LLM, yalnızca yapılandırılmamış verileri yorumlamak için bir anlamsal işlemci olarak kullanılmaktadır. Karar verme yetkisi (özellikle risk puanlama mantığı) ise ayrıştırılarak şeffaf ve deterministik bir Metrik Motoru tarafından yönetilmektedir. Böylece, risk değerlendirmesini güvenli bir şekilde otomatize etmek için Üretken Yapay Zeka, titiz ve uzmanlarca doğrulanmış koruma sınırları ile sınırlandırılmaktadır.

Bu çalışma, kesin olarak tanımlanmış bir metrik mimarisi içinde kamuya açık LLM'lerin (özellikle GPT-4o, Gemini 2.0 Flash) akıl yürütme yeteneklerinden yararlanan, otomatik ve uzman onaylı bir Risk Değerlendirme Çerçevesi sunmaktadır. Yapay zekadan kısıtlanmamış bir şekilde "risk değerlendirmesi" yapmasını istemek yerine, akıl yürütme süreci hassas bir "Dinamik Metrik Motoruna" dayandırılmaktadır. Bu motorun parametreleri ve ağırlıkları rastgele seçilmemiş; 101 siber güvenlik profesyoneli ile yapılan kapsamlı bir anketten ROC yöntemi kullanılarak türetilmiştir. Bu yaklaşımda veri analizinin ağır yükünü yapay zeka üstlenirken, karar mantığının ampirik olarak elde edilen uzman yargılarına dayalı kalmasını sağlamaktadır.

Önerilen çerçeve, mevcut literatüre şu katkıları sağlamaktadır:

- **Uzman Onaylı Metrik Tasarımı:** 101 endüstri uzmanının sıra tabanlı derecelendirmelerini kardinal ağırlıklara dönüştürmek için ROC yöntemini kullanan, ampirik verilere dayalı bir risk puanlama modeli sunulmuştur.
- **LLM Destekli Otomasyon:** Ticari LLM'lerin otomatik risk analistleri olarak nasıl yapılandırılabileceği; heterojen girdileri (CTI raporları, varlık profilleri) ayrıştırarak bunları yüksek sadakatle standart bir metrik şemasına nasıl eşleştirebileceği açıklanmıştır.
- **Karşılaştırmalı Doğrulama (İnsan ve YZ):** On kıdemli uzman ve iki LLM ajanı (GPT-4o ve Gemini 2.0 Flash) tarafından değerlendirilen on beş gerçek dünya senaryosu (C_1--C_{15}) aracılığıyla, değerlendirme döngü süresini 100 kattan fazla iyileştirirken insan medyanıyla yakın bir uyum ($\sigma \approx$ 0,15-0,25) sağlandığına dair ampirik kanıtlar sunulmuştur.

Bu makalenin geri kalanı şu şekilde düzenlenmiştir: Bölüm 2, ilgili çalışmaları incelemekte ve mevcut metodolojilerin sınırlamalarını tartışmaktadır. Bölüm 3; metodolojiyi, özellikle anket tasarımını ve ROC ağırlık türetimini detaylandırmaktadır. Bölüm 4, önerilen çerçeve mimarisini ve YZ ajanı iş akışını tanıtmaktadır. Bölüm 5, karşılaştırmalı vaka çalışmasını ve değerlendirme sonuçlarını sunmaktadır. Bölüm 6, algoritmik yanlılık ve sınırlamalar dahil olmak üzere modellerin davranışsal özelliklerini analiz etmektedir. Sonuç ve Öneriler bölümü ile çalışmayı sonuçlandırmakta ve gelecekteki araştırma yönlerini tartışmaktadır.

BÖLÜM 2

2. LİTERATÜR ARAŞTIRMASI

Siber güvenlik risk değerlendirmesine, yönetişim odaklı kılavuzlar, standartlar ve puanlama yapılabilmesine olanak sağlayan çerçeveler odağında bağlamsal ve yapay zeka odaklı unsurların eklenmesine yönelik çeşitli çalışmalar akademik çizgide geliştirilmiştir. Bu bölümde, yaklaşımları gözden geçirip, sınırlamalarını özetlenmekte ve literatürdeki boşluklar belirlenmektedir.

2.1 STANDART PUANLAMA METODOLOJİLERİ

Güvenlik açıklıklarının puanlanmasına olanak sunan CVSS ve risk faktörü analizinde faydalanılan FAIR gibi nicel puanlama sistemleri yaygın olarak benimsenmektedir (Forum of Incident Response and Security Teams [FIRST], 2023; Freund ve Jones, 2014). CVSS' in v2, v3.1 ve v4.0 sürümlerinde saldırı vektörü, ayrıcalık gereksinimi ve kullanıcı etkileşimi gibi parametreler kullanılarak, açıklıkların puanlanması standartlaştırılır. FAIR, finansal açıdan riski tahmin etmek için sıklık ve büyüklüğe odaklanırken, Ortak Zayıflık Puanlama Sistemi (CWSS) değerlendirmeyi yazılım kusurlarına da genişletir (Forum of Incident Response and Security Teams [FIRST], 2023). Bu modeller, niteliksel yaklaşımlara nazaran daha tekrarlanabilir sonuçlar üretse de, esasen statik kalmaktadır. Süreçlerine CTI dahil edilemez, organizasyonel bağlam ihmal edilir ve ağırlıklandırma yapıları uyarlanabilir değil, ön tanımlıdır.

Tablo 2.1’ de görüleceği üzere, mevcut akademik modeller İstismar Edilebilirlik ve Etki gibi geleneksel boyutları kapsamakta. Buna rağmen NLP/LLM entegrasyonu ve kurumsal bağlam farkındalığı konularında sınırlı bir kapsam sergilemektedir. Özellikle, çalışmaların yalnızca küçük bir azınlığı CTI akışlarını doğrudan süreçlerine dahil etmekte ve sınırlı sayıda çalışma otomatik

muhakeme için YZ yararlanmaktadır. bu durum, önerilen çerçevenin ele almayı amaçladığı boşluğu göstermektedir.

Tablo 2.1 Akademik Çalışmalarının Karşılaştırmalı Özeti

Çalışma	Kültürel Boyutlar	ÖMA	CTI	NLP/ LLM	YZ/ MÖ
(Kawanishi vd., 2023)	SÖM, SAL, ETK, İŞ, BAĞ	✓	X	X	X
(Wang, W. vd., 2020)	SÖM, ETK, BAĞ	✓	X	X	X
(Wang, T. vd., 2020)	SÖM, SAL, ETK, BAĞ	✓	X	X	X
(Ahmed vd., 2022)	SÖM, SAL, ETK, İŞ, BAĞ, KUR	✓	✓	X	X
(Yang vd., 2023)	SÖM, SAL, ETK, BAĞ	✓	X	X	X
(Aksu vd., 2017)	SÖM, SAL, ETK, BAĞ	✓	X	X	X
(Younang ve Sen, 2025)	SÖM, SAL, ETK, BAĞ	X	X	X	X
(Bansal vd., 2024)	SÖM, SAL, ETK, BAĞ	✓	X	X	X
(Abbas vd., 2025)	SÖM, SAL, ETK, BAĞ	X	✓	X	✓
(Moreira vd., 2021)	ETK, İŞ, BAĞ, KUR	✓	X	X	X
Önerilen Çerçeve	SÖM, SAL, ETK, İŞ, BAĞ, KUR	✓	✓	✓	✓

Açıklamalar – Geleneksel ve Bağlamsal Risk Boyutları: SÖM: Sömürülebilirlik, **SAL:** Saldırgan Modelleme, **ETK:** Teknik Etki, **İŞ:** İş/Mevzuat Etkisi, **BAĞ:** Varlık/Hedef Bağlam Farkındalığı, **KUR:** Kurumsal Faktörler.

Açıklamalar – Yapay Zeka Destekli ve Veri Odaklı Yetenekler: ÖMA: Özelleştirilebilir Metrik Ağırlıklandırma, **CTI:** Tehdit İstihbaratı Entegrasyonu, **NLP/LLM:** Doğal Dil İşleme / Büyük Dil Modelleri, **YZ/MÖ:** Yapay Zeka / Makine Öğrenmesi.

2.2 GELENEKSEL RİSK DEĞERLENDİRME METODOLOJİLERİ

NIST SP 800-30 (National Institute of Standards and Technology [NIST], 2012), ISO/IEC 27005 (International Organization for Standardization [ISO], 2022), ISO 31000 (International Organization for Standardization [ISO], 2018)

ve OCTAVE (Alberts ve Dorofee, 2002) gibi metodolojiler risklerin tanımlanması, analiz edilmesi ve değerlendirilmesi için kılavuzlar sunmaktadır. Bu metodolojiler; uyumluluk ve yönetişim alanlarında temel teşkil etmekte, güvenlik ve risk yönetimi programları için yaygın olarak tanınan standartlar olarak hizmet vermektedir. Ancak, uzman görüşüne olan bağımlılıkları bu yöntemleri büyük ölçüde öznel kılmakta; bu durum da organizasyonlar arasında tutarsızlıklara ve sınırlı tekrarlanabilirliğe yol açmaktadır. 5 X 5 risk matrisleri gibi araçlar olasılık ve etkinin nitel olarak tanımlamasın olanak sunarken, riskleri şeffaf alt metrikler olmaksızın yüksek, orta ve düşük risk kategorileriyle eşleştirilmektedir. Bu tür çıktılar üst düzey raporlama için değerli olsa da hızla gelişen tehdit ortamlarında dinamik ve veri odaklı karar verme süreçleri için yetersiz kalmaktadır.

Tablo 2.2' de yaygın olarak kullanılan çerçevelerin karşılaştırmalı bir özetini sunmakta. Önerilen modelin bu çerçeveler arasındaki konumunu görülebilir.

2.3 BAĞLAM TEMELLİ VE GÜNCEL YAKLAŞIMLAR

Güncel araştırmalar; saldırgan modelleme, bağımlılık analizi ve iş bağlamını sürece dahil ederek zafiyet puanlamasını zenginleştirmeyi amaçlamakta. Zamana duyarlı olasılıksal risk değerlendirme modelleri (Cheimonidis ve Rantos, 2025) ve MITRE ATT&CK tabanlı saldırgan davranış modellemesi (Al-Sada vd., 2024) bu duruma örnek teşkil etmektedir. Bu ilerlemelere rağmen zamansal dinamikler veya saldırgan davranışı gibi belirli boşluklar ele alınmakla birlikte; tehdit istihbaratı, kurumsal bağlam, otomatik hesaplama ve döngüde insan (HITL) yönetişimini birleştiren bütünleşik bir mimari sunulamamaktadır.

Örneğin Wang vd. (2020), Sağlık Nesnelerin İnterneti (IoHT) cihazlarını önceden tanımlanmış güvenlik niteliklerine göre değerlendirmek ve sıralamak üzerine çalışnmıştır. Bu amaçla hibrit AHP-TOPSIS yöntemine dayalı bir ISA değerlendirme çerçevesi önermektedir.

Tablo 2.2 Risk Değerlendirme Çerçevelerine Karşılaştırmalı Genel Bakış

Kriter / Boyut	NIST SP 800-30 (NIST, 2012)	FAIR (Freund ve Jones, 2014)	OCTAVE (Alberts ve Dorofee, 2002)	Önerilen Çerçeve
Kapsam ve Amaç	Tanımlama-Analiz-Değerlendirme	Nicel (frekans × büyüklük)	Senaryo tabanlı tanımlama	Tam yaşam döngüsü (Otomatik)
Hesaplama Yaklaşımı	Nitel/Yarı-nicel O × E	Frekans × Büyüklük	Nitel senaryolar	LLM Veri Çıkarımı → ROC
Temel Parametreler	Tehdit/zafiyet/etki	Aktör/varlık/kontroller	Varlık/süreç/tehditler	30+ metrik + bağlam
Bağlam Farkındalığı	Orta	Yüksek	Orta-Yüksek	Çok Yüksek (LLM Destekli)
Saldırgan Modelleme	Kısıtlı	Kısmi	Yok	Detaylı (13+ metrik)
Kurumsal Faktörler	Kısmi	Kısıtlı	Yüksek (süreç odaklı)	Tam entegre
CTI Entegrasyonu	Yok	Dolaylı	Yok	Tam entegrasyon
YZ/MÖ Yetenekleri	Yok	Yok	Yok	Üretken Yapay Zeka (GPT-4o)
NLP/LLM Yetenekleri	Yok	Yok	Yok	Temel Motor
Uyarlanabilirlik	Yok	Yok	Yok	Bağlama Duyarlı (LLM)
Geri Bildirim Mek.	Yok	Yok	Yok	HITL geri bildirim döngüsü
Otomasyon Seviyesi	Manuel	Kısmi	Manuel	Yüksek (>100 kat hızlanma)
Özelleştirilebilirlik	Düşük-Orta	Yüksek	Yüksek	Çok Yüksek
İyileştirme Desteği	Dolaylı	Dolaylı	Dolaylı	Yerleşik

Söz konusu çalışma, çok kriterli karar verme tekniklerinin sağlık hizmetlerine özel bir IoT bağlamında uygulanabilirliğini gösterse de büyük ölçüde statik ve alan odaklıdır. Buna karşılık mevcut çerçeve;

(i) anket verilerinden türetilen bir ROC ağırlıklandırma şemasına dayalı LLM destekli bağlamsal çıkarım,

(ii) periyodik yeniden kalibrasyon içeren bir HITL geri bildirim döngüsü,

(iii) dinamik siber güvenlik risk değerlendirmesini desteklemek için alanlar arası bağlamsal veri füzyon,

entegre ederek, çalışmasını IoT'ye özel ortamların ötesine genelleştirmektedir.

Ayrıca, CTI paylaşımı alanında süreci işbirlikçi oyun teorisi ve Shapley değeri kullanarak modelleyen yaklaşımlar da ortaya çıkmıştır. Bu yaklaşımlarda fayda dağıtım mekanizması, dirençli ve adil sonuçları teşvik etmek amacıyla bağlamsal parametrelerle (örneğin bir risk katsayısı) kullanılmaktadır (Xie vd., 2025).

2025 yılında literatür, siber güvenlik tespit yeteneklerini geliştirmek için Makine Öğrenmesi (ML) tekniklerinin entegrasyonuna yönelik artan bir vurguya işaret etmektedir. Femi ve Madu (2025), yapay zeka destekli güvenlik sistemlerinin vaka oranlarını azaltabildiğini bildirmiş ve fidye yazılımlarının %98'e varan oranla engellenebildiğini vurgulamışlardır. Bu tür yeteneklerin bulut ortamlarında işlevsel hale getirmek için Jamili vd. (2025), Rastgele Orman (Random Forest) sınıflandırıcılarını otokodlayıcı tabanlı anomali tespitiyle birleştiren hibrit bir çerçeve önermiş ve %95,3'lük bir doğruluk oranına ulaşıldığını bildirmişlerdir. Bununla birlikte yaklaşımlar; gerçek zamanlı kurumsal ortamlarda elde edilmesi genellikle zor olan, büyük miktarda etiketlenmiş veri gerektiren denetimli öğrenme bileşenlerine ihtiyaç duymaktadır.

ML tabanlı yaklaşımlardaki temel zorluklardan biri, modellerin "kara kutu" (black-box) doğasıdır. Islam vd. (2025), sinir ağlarını SHAP ve LIME gibi Açıklanabilir Yapay Zeka (XAI - Explainable AI) teknikleriyle birleştirerek bu sorunu ele almaya çalışmıştır. Hesaplama karmaşıklığının artması pahasına yorumlanabilirliği iyileştirebilmişlerdir. Benzer şekilde Malik vd. (2025), Yapay Sinir Ağlarını (ANN) Yorumlayıcı Yapısal Modelleme (ISM) ile birlikte kullanmışlardır. Ancak nesnel metrikler yerine öznel uzman anketlerine

dayanmaları, eğitim sürecine potansiyel yanlılığın dahil etmeklerine sebep olmuştur. Simülasyon tarafında ise Camacho vd. (2025), Çekişmeli Risk Analizi'ne (ARA - Adversarial Risk Analysis) dayalı bir çerçeve önermektedir. Yaklaşım finansal risklerin azaltılmasında etkili olsa da önemli ölçüde modelleme çabası ve kapsamlı alan uzmanı katılımına ihtiyaç duymakta, bu da operasyonların otomatikleştirilmesi için ölçeklenebilirliğini sınırlamaktadır.

Siber tehdit çıkarımı ve bağlamsal profillemenin otomatize edilmesinde Doğal Dil İşleme (NLP) ve LLM tabanlı modellerin etkinliğini ortaya koyacak güncel çalışmalar bulunmaktadır. Örneğin Hmimou vd. (2025), yapılandırılmamış veri kaynakları üzerinde tehdit varlıklarını anlamsal olarak ilişkilendiren çok ajanlı bir LLM çerçevesi sunarken; Marinho ve Holanda (2023), yeni ortaya çıkan tehditleri tanımlamak ve bağlamsal profiller oluşturmak amacıyla NLP destekli sınıflandırma ve varlık tanıma tekniklerini uygulamaktadır. Söz konusu bu yaklaşımlar çalışmada önerilen LLM katmanının amaçlarıyla paralellik göstermektedir.

2.4 LİTERATÜRDEKİ EKSİKLİKLER VE ÖZGÜN KATKI

Gerçekleştirilen literatür taraması neticesinde iki temel eksiklik ortaya çıkmaktadır. Birincisi, mevcut yaklaşımların çoğu ölçeklenebilirlik-hassasiyet ikilemi (scalability-precision dilemma) ile karşı karşıyadır. Bu durum ya statik ve hızlı ancak bağlamdan bağımsızdırlar (örneğin puanlama sistemleri) ya da bağlama duyarlı ancak emek yoğun ve yavaştırlar (örneğin nitel çerçeveler) (Zhang vd., 2023).

İkinci olarak, siber güvenlikte LLM'lerin potansiyeli kabul edilse de bunların nicel risk değerlendirmesine entegrasyonu büyük ölçüde keşif aşamasında kalması. Mevcut literatür, halüsinasyonları azaltmak amacıyla Üretken Yapay Zekayı ampirik uzman bilgisiyle nasıl sınırlandırılacağına dair sınırlı metodolojik rehberlik sunmaktadır. Dahası, güncel ML tabanlı çalışmalarda gözlemlendiği üzere (Islam vd., 2025; Jamili vd., 2025), geleneksel modeller sıklıkla "kara kutu" davranışı sergilemekte ya da kapsamlı eğitim veri

setlerine ihtiyaç duymaktadır. Simülasyon tabanlı çerçevelerde ise (Camacho vd., 2025) önemli ölçüde modelleme çabasına ihtiyaç duyulmaktadır. Geliştirilen çerçeve;

(i) çok kaynaklı veri girişi (multi-source data ingestion),

(ii) kamuya açık LLM (GPT-4o ve Gemini 2.0 Flash) aracılığıyla otomatik muhakeme ve

(iii) şeffaf, anket temelli bir metrik motorunu birleştirerek literatürdeki boşlukları gidermek üzere tasarlanmıştır.

Kara kutu (black-box) yaklaşımlarının aksine, yapay zeka destekli analizin uzmanlarca doğrulanmış ağırlıklara sabitlenebilmesi için ROC yöntemi kullanılmıştır. Tablo 2.1 ve Tablo 2.2'de özetlendiği üzereönerilen model hem otomatik hem de uzman onaylı bir yaklaşım sunarak statik puanlama ile manuel değerlendirme arasındaki boşluğu kapatmayı hedeflemektedir.

BÖLÜM 3

3. UZMAN GÖRÜŞÜNE DAYALI METRİK TASARIMI

Bu bölüm, önerilen çerçevenin uzman görüşüne dayalı nicelleme modelini (expert-informed quantification model) açıklayarak çalışmanın metodolojik temelini sunmaktadır. Şeffaf olmayan karar mantığına sahip "kara kutu" yapay zeka yaklaşımları benimsemek yerine, çalışma kapsamında şeffaf ve anket temelli bir metrik motoru kullanılmaktadır. Bölüm 4' de sistem tasarımı sunulurken; bu bölümde tekrarlanabilirlik, açıklanabilirlik ve statik puanlama yaklaşımlarına kıyasla sistematik iyileştirmeyi desteklemek amacıyla ağırlık tespit edilmesi ve parametrelerin kalibrasyon metodolojisi detaylandırılmaktadır.

3.1 PARAMETRE VE METRİK AĞIRLIKLARININ BELİRLENMESİ İÇİN ANKET TASARIMI

3.1.1 Anketin Amacı

Anketin temel amacı, risk puanlama metodolojisini öznel ve sezgisel (heuristic-based) bir yaklaşımdan daha nesnel ve veri odaklı bir çerçeveye dönüştürmektir. Bu doğrultuda anket tasarımında şu hedefleri gözetilmiştir:

- Çerçevede tanımlanan risk parametreleri için ampirik önem sıralamaları elde ederek ROC yöntemi aracılığıyla temel ağırlıkların hesaplamak (Bkz: Tablo 3.1 , Tablo 3.2),
- Risk hesaplama bileşeni içinde tutarlılığı teşvik etmek amacıyla alt metriklerin şiddet puanlarını standart bir ölçekte kalibre etmek,
- Tek bir analistin yargısına olan bağımlılığı azaltmak ve daha tutarlı değerlendirmeleri desteklemek,

- Önerilen çerçevenin ilk uygulaması için uzman görüşüne dayalı bir temel (baseline) oluşturmak.

Tablo 3.1 Türetilen Ağırlıklarla Risk Hesaplama Metrikleri Kutusu

Temel Parametreler	Puanlı Alt Parametreler	WF	(O/E)
Saldırı Yüzeyi ve Sömürülebilirlik			
Saldırı Vektörü	Fiziksel (0,29), Yerel (0,56), Yakın Ağ (0,78), Ağ (0,99)	0,1341	O
Gereken Yetkiler	Yüksek (0,39), Düşük (0,91), Yok (1,00)	0,0665	O
Kullanıcı Etkileşimi	Gerekli (0,65), Yok (1,00)	0,0437	O
Sömürü Zinciri	Kontrollü (0,53), Kontrolsüz (0,91)	0,0121	O
Kimlik Doğrulama Karmaşıklığı	Çok Faktörlü (0,17), İki Faktörlü (0,47), Tek Faktörlü (0,99), Yok (1,00)	0,0902	O
Yama Durumu	Tam Yamalı (0,25), Kısmi Yama (0,88), Yama Yok (1,00)	0,1078	O
Maruziyet Süresi	0-30 Gün (0,63), 1-6 Ay (0,80), 6+ Ay (0,95)	0,0057	O
Sömürü Olgunluğu	Mümkün (0,72), Aktif (1,00)	0,0193	O
Tehdit Aktörü ve Yetenekleri			
Gereken Yetenekler	Yetkin (0,31), Uzman (0,50), Operasyonel (0,74), Minimum (0,91), Yok (1,00)	0,0503	O
Kaynaklar	Bireysel (0,40), Ekip (0,69), Kurumsal (0,89), Devlet (0,90)	0,1867	O
Hedefler	Kopyalama (0,69), Hasar (0,78), Yok Etme/Engelleme (0,82), Ele Geçirme (0,85), Hepsi (0,98)	0,0378	O
Sınırlar	Davranış Kuralları (0,22), Yasal (0,50), Yasa Dışı (0,91)	0,0028	O
Görünürlük	Açık (0,39), Gizli (0,87)	0,0088	O
Otomasyon Seviyesi	Manuel (0,44), Betik (0,87), Tam Otomatik (0,95)	0,0326	O
Motivasyon Türü	Düşman Olmayan (0,18), Aktivizm (0,63), İç Tehdit (0,75), Casusluk (0,76), Finansal (0,84)	0,0771	O
Tedarik Zinciri Saldırısı	Hayır (0,59), Evet (0,87)	0,0278	O

Tablo 3.1 (Devamı) Türetilen Ağırlıklarla Risk Hesaplama Metrikleri Kutusu

Temel Parametreler	Puanlı Alt Parametreler	WF	(O/E)
Hedef Profili ve Maruziyet			
Marka Değeri	Düşük (0,21), Orta (0,49), Yüksek (0,77), Kritik (0,93)	0,0234	O
Çalışan Farkındalığı	İleri (0,24), Orta (0,57), Temel (0,94), Yok (1,00)	0,0156	O
Olgunluk Seviyesi	Optimize (0,12), Yönetilen (0,31), Tanımlı (0,52), Gelişmekte (0,78), Başlangıç (0,97)	0,0578	O
Sektörel Hassasiyet	Düşük (0,20), Orta (0,53), Yüksek (0,83), Kritik (0,95)	0,1753	E
Olay Müdahale Kabiliyeti	İleri (0,20), Yeterli (0,50), Yavaş (0,81), Mevcut Değil (0,97)	0,0425	E
Mevzuat Etkisi	Yok (0,10), Düşük (0,36), Orta (0,67), Sıkı (0,95)	0,1058	E
Hasar ve Etki			
Gizlilik Etkisi	Yok (0,00), Düşük (0,28), Orta (0,67), Yüksek (0,95)	0,2586	E
Bütünlük Etkisi	Yok (0,00), Düşük (0,27), Orta (0,65), Yüksek (0,91)	0,0544	E
Erişilebilirlik Etkisi	Yok (0,00), Düşük (0,33), Orta (0,67), Yüksek (0,92)	0,1336	E
Teknik Etki	Etki Yok (0,00), Minimum (0,22), Orta (0,51), Şiddetli (0,78), Kritik (0,93)	0,0321	E
Etki Alanı	Tekil (0,30), Çoklu (0,71), Yaygın (0,94)	0,0229	E
Finansal Etki	Yok (0,00), Küçük (0,32), Büyük (0,69), Felaket (0,89)	0,0683	E
İtibar Etkisi	Yok (0,11), Düşük (0,40), Yüksek (0,80), Onarılamaz (0,94)	0,085	E
Süre	Geçici (Saat) (0,26), Geçici (Gün) (0,52), Geçici (Ay) (0,77), Kalıcı (0,96)	0,0069	E
İş Sağlığı ve Güvenliği Etkisi	Yok (0,00), Küçük (0,27), Büyük (0,55), Aşırı (0,74)	0,0145	E

Tablo 3.2 Temel Parametre Listesi ve Tanımları

Temel Parametreler	Tanım (Açıklama)
Saldırı Vektörü	Fiziksel, yerel veya ağ üzerinden erişim gereksinimleri.
Gereken Yetkiler	Saldırının gerçekleştirilmesi için gereken kullanıcı yetki düzeyi.
Kullanıcı Etkileşimi	Saldırının başarılı olması için bir kullanıcının eylemine ihtiyaç duyulup duyulmadığı.
Sömürü Zinciri	Zafiyetin bir zincirleme saldırı içerisinde kullanılıp kullanılamayacağı.
Kimlik Doğrulama Karmaşıklığı	Sömürü için gereken ek kimlik doğrulama katmanları.
Yama Durumu	Zafiyet için mevcut bir yamanın olup olmadığı.
Maruziyet Süresi	Zafiyetin ne kadar süredir bilindiği veya istismar edildiği.
Sömürü Olgunluğu	Sömürü kodunun veya yönteminin kullanılabilirlik düzeyi.
Gereken Yetenekler	Saldırganın sahip olması gereken teknik beceri düzeyi.
Kaynaklar	Saldırı için gereken maddi veya operasyonel kaynaklar.
Hedefler	Hedeflenen hasar veya veri sızdırma seviyesi.
Sınırlar	Saldırganın yasal, teknik veya etik kısıtlamaları.
Görünürlük	Saldırının tespit edilme olasılığı.
Otomasyon Seviyesi	Saldırının otomatik araçlarla yürütülme derecesi.
Motivasyon Türü	Saldırganın birincil amacı (finansal, casusluk vb.).
Tedarik Zinciri Saldırısı	Tedarik zinciri veya "ada sıçraması" tekniklerinin kullanımı.
Marka Değeri	Kuruluşun itibarının ve marka algısının önemi.
Çalışan Farkındalığı	Çalışanların siber güvenlik bilinç düzeyi.
Olgunluk Seviyesi	Kuruluşun siber güvenlik süreç olgunluk düzeyi.
Sektörel Hassasiyet	Sektörün ulusal veya uluslararası güvenlik açısından kritikliği.
Olay Müdahale Kabiliyeti	Tehditleri tespit etme ve hafifletme hızı.

Tablo 3.2 (Devamı) Temel Parametre Listesi ve Tanımları

Temel Parametreler	Tanım (Açıklama)
Mevzuat Etkisi	Uyum zorunlulukları ve düzenleyici denetim düzeyi.
Gizlilik Etkisi	Veri gizliliği üzerindeki potansiyel hasar.
Bütünlük Etkisi	Verilerin manipüle edilme veya değiştirilme riski.
Erişilebilirlik Etkisi	Sistem kesintisi veya hizmet dışı kalma riski.
Teknik Etki	Teknolojik altyapıda meydana gelen hasar.
Etki Alanı	Etkilenen sistemlerin kapsamı (tekil veya yaygın).
Finansal Etki	İhlalden kaynaklanan tahmini parasal kayıp.
İtibar Etkisi	İhlalin kamuoyuna duyurulması sonucu markaya verilen zarar.
Süre	Etkinin kalıcılığı (saatler, günler veya aylar).
İş Sağlığı ve Güvenliği Etkisi	İnsan sağlığı ve fiziksel güvenlik üzerindeki etkisi.

3.1.2 Anket Yapısı ve Soru Türleri

Uzman bilgisinin farklı boyutlarını kapsamak amacıyla anket üç ana bölümden oluşacak şekilde tasarlanmıştır (Bkz: Şekil 3.1).

- **Katılımcı Profili:** Uzmanlık seviyelerini tanımlamak amacıyla katılımcıların organizasyondaki görevi (örn. CISO, analist), deneyim sürelerine ve organizasyonun faaliyet gösterdikleri sektöre ilişkin veriler toplanmıştır.
- **Parametre Sıralaması (ROC için):** Katılımcılardan Olasılık ve Etki gruplarına ait parametreleri "En Kritik"ten "En Az Kritik"e doğru sıralamaları istenmiştir. Elde edilen bu veriler, ROC ağırlık türetimi yöntemi için doğrudan girdi olarak kullanılmıştır.
- **Metrik Puanlama (0 ile 10 arası):** Katılımcılardan parametrelere ait alt parametrelere (örn. "İstismara Hazır Zafiyet - Exploit Available" ile

"Yama Mevcut Değil - No Patch" kıyaslaması) 0 ile 10 arasında sayısal değerler tanımlanması istenmiştir.

Demografi, parametre sıralaması (ROC) ve metrik kalibrasyonu için veri toplama stratejisini gösteren anket aracından örnek sorular.

Anket aracı, Metrik Motoru için uzman bilgisini toplamak amacıyla tasarlanmış üç ayrı bölümden oluşmaktadır.
Tip 1: Demografi ve Uzmanlık Profili *S:* "Aşağıdakilerden hangisi siber güvenlikteki birincil rolünüzü en iyi şekilde tanımlar?"

- [ ] CISO / Yönetici
- [ ] Güvenlik Mimarı / Mühendisi
- [ ] Risk Analisti / Danışmanı
- [ ] SOC Analisti / Olay Müdahale Uzmanı

Tip 2: Parametre Sıralaması (ROC Ağırlıkları İçin) *S:* "Lütfen aşağıdaki **Tehdit Aktörü Yeteneklerini**, başarılı bir ihlal **olasılığına** katkılarına göre 'En Kritik' (1) ile 'En Az Kritik' (6) arasında sıralayınız."

1. Kaynaklar (Zaman, Finansman)
2. Teknik Beceriler
3. Motivasyon
4. Erişim Hakları (Kurum İçi) *(Arayüz: Sürükle-bırak yöntemiyle sıralanabilir liste)*

Tip 3: Metrik Kalibrasyonu (0-10 Ölçeği) *S:* "0 ile 10 arasındaki bir ölçekte, aşağıdaki **Saldırı Vektörü** koşullarının şiddetini nasıl puanlarsınız?"

- **Ağ:** Uzaktan sömürülebilir. *[Kaydırıcı: 0-10]*
- **Yerel:** Kabuk (shell) erişimi gerektirir. *[Kaydırıcı: 0-10]*
- **Fiziksel:** Fiziksel etkileşim gerektirir. *[Kaydırıcı: 0-10]*

Şekil 3.1 Uzman Anketin Örnek Sorular

3.2 PARAMETRE GRUPLARI VE SIRALAMA DÜZENİ

Önceki çalışmaya dayalı olarak (Unal ve Celiktas, 2025), siber riskin çok boyutlu doğasını temsil etmek amacıyla "Metrik Kutusu" olarak adlandırılan yüksek ayrıntılı bir parametre/altparametre yapısı kullanılmaktadır. Buna göre anket tasarımı, Olasılık ve Etki olmak üzere iki ana boyuta ayrılmıştır. ROC yöntemi aracılığıyla ağırlıkların türetilmesi için kullanılacak sıralamalarda uzman yargılarında tutarlılığı teşvik etmek amacıyla hiyerarşik bir sıralama

stratejisi tercih edilmiştir. Bu süreçte izlenen adımlar aşağıda detaylandırılmaktadır:

- **Bilişsel Hazırlık (Grup İçi Sıralama):** Katılımcılar ilk olarak mantıksal alt gruplar (4 temel grup; örn. Saldırı Yüzeyi, Tehdit Aktörü Yetenekleri) içindeki parametreleri önceliklendirmeleri istenmiştir. Bu adım, uzmanların daha karmaşık olan genel sıralamayı gerçekleştirmeden önce, ilgili faktörlerin göreceli önemini izole bir şekilde değerlendirmelerine olanak tanıyan bir hazırlık görevi görmüştür.
- **Genel Sıralama:** Hazırlık aşamasının ardından katılımcılar, Olasılık boyutundaki tüm parametreleri (19 madde) ve Etki boyutundaki tüm parametreleri (12 madde) iki ayrı soruda sıralamışlardır. Elde edilen sonuçlar, her uzmanın risk ortamına ilişkin bütünsel değerlendirmesini yansıtmaktadır. Bu nedenle ROC ağırlık hesaplaması için temel girdi olarak kullanılmıştır.

3.3 METRİK (ALT PARAMETRE) PUANLAMA YÖNTEMİ VE MANTIĞI

Sıralama süreci her bir parametrenin ağırlığını belirlerken, nicel risk hesaplaması için her bir parametreye ait alt parametre durumu için kalibre edilmiş şiddet eğerlerine ihtiyaç duyulmaktadır. Bu değerlerin elde edilmesi amacıyla uzmanlar, ankette sunulan kalibrasyon kurallarına dayanarak 0 ile 10 ölçeğinde puan atamaları gerçekleştirmişlerdir. Söz konusu puanlama mantığı şu esaslara dayanmaktadır:

- **Taban Çizgisi (Puan 0):** Riskin, şemada tanımlanan en düşük seviyeye indirildiği durumu temsil etmektedir (örneğin, Yama Mevcudiyeti: Tamamen Yamalanmış). Uzmanlara, bir durumun ilgili Olasılık veya Etki katkısını ortadan kaldırdığı durumlarda 0 puan atamaları yönünde rehberlik edilmiştir.

- **Tavan (Puan 10):** Teorik en kötü durumu, yani ilgili parametre için maksimum risk katkısını temsil etmektedir (örneğin, Yama Mevcudiyeti: Yama Yok / Sıfırıncı Gün).
- **Aralık Kalibrasyonu:** Ara puanlar, bu iki uç nokta arasındaki göreceli risk büyüklüğünü yansıtmaktadır. Bu yaklaşım, çerçevenin ikili (binary) girdilere güvenmek yerine; granüler farkları (örneğin, "Yerel Erişim" ile "Ağ Erişimi" arasındaki şiddet farkı) yakalamasına olanak tanımaktadır.

3.4 VERİ TOPLAMA VE ANKET UYGULAMA

Anket süreci, arayüz yetenekleri ve sahip olduğu güvenlik özellikleri nedeniyle tercih edilen Jotform çevrimiçi anket platformu üzerinden yürütülmüştür. Hazırlanan anket, hedef uzman grubuna doğrudan bir erişim bağlantısı gönderimi ile dağıtılmıştır. Veri kalitesini desteklemek ve sıralama görevleri sırasında bilişsel yükü azaltmak amacıyla, sürükle-bırak arayüzüne sahip özel bir "Sıralanabilir Liste" (Orderable List) eklentisi tercih edilmiştir. Tercih edilen bu eklenti aşağıda belirtilen avantajları sağlamıştır:

- **Bütünsel Görünürlük:** Katılımcıların tüm parametre setini tek bir ekranda eş zamanlı olarak görmelerine olanak tanınmıştır. Burada çok sayfalı veya açılır menü tabanlı sıralama formatlarında sıklıkla görülen bellek kaybı ve dikkat dağılması en aza indirgenmiştir.
- **Dinamik ve Hızlı Karşılaştırma:** Sürükle-bırak arayüz, uzmanların öğeleri yukarı-aşağı hareket ettirerek öncelikleri anlık olarak ayarlamasına imkan sağlamıştır. Tüm parametrelerin sunulduğu liste içerisinde ikili karşılaştırmalara imkan sağlamış, birbirine yakın parametreler arasındaki karşılaştırmaları kolaylaştırarak katılımcıların hiyerarşiyi kendi profesyonel yargılarını tam olarak yansıtana kadar iyileştirmelerine zemin hazırlamıştır.

3.5 AĞIRLIK TÜRETME YAKLAŞIMI

Uzman görüşlerini puanlama modeline dönüştürmek amacıyla;

- parametre ağırlıkları için ROC (Rank Order Centroid) ve
- alt parametre puanları için normalizasyon ve sabitleme (normalization-and-anchoring) tekniğinden oluşan çift katmanlı bir kalibrasyon yaklaşımı kullanılmıştır.

3.5.1 ROC Yöntemi ile Parametre Ağırlıklandırma (WF)

Ana parametrelerin (örn. Saldırı Vektörü, Gizlilik Etkisi) ağırlık değerleri (*WF*), ROC yönteminden faydalanılarak türetilmiştir. Bu yaklaşım, uzmanlardan toplanan öncelik sıralı parametreleri nicel ağırlıklara dönüştürmekte. Bu yöntem puanlama ile ilişkili yanlılıkları en aza indirgenmesinde yaygın olarak faydalanılmaktadır. *M* sayıdaki öğe arasında *k* en önemli öğe için ROC ağırlığı (w_k) şu şekilde hesaplanır:

$$w_k = \frac{1}{M}\sum_{i=k}^{M}\frac{1}{i} \tag{3.1}$$

Tekrarlanabilirliğin sağlanması amacıyla parametreler Olasılık (*O*) ve Etki (*E*) olmak üzere iki boyutta ele alanmaktadır. Her bir parametrenin hangi boyutun hesaplamasında kullanılıdığı Tablo 3.1 üzerinde gösterilmiştir. ROC yöntemiyle türetilen ağırlıklar (w_k), her bir boyut için ağırlıklar toplamı bire ($\sum_{j\in O} w_j \approx 1.0$ ve $(\sum_{j\in E} w_j \approx 1.0)$) eşit olacak şekilde, kendi boyutları dahilinde bağımsız olarak normalize edilmiştir.

3.5.2 Deterministik Risk Puanı Hesaplaması

$s_j \in [0,1]$ ifadesi, *j* metriği için seçilen normalize edilmiş şiddet puanını; w_j ise ilgili ROC yöntemiyle türetilmiş ağırlığı temsil etmektedir. Burada Olasılık, başarılı bir istismarın (exploitation) gerçekleş olasılığını ifade etmektedir. Olasılık (*O*) ve Etki (*E*) bileşenleri şu formüller yardımıyla hesaplanmaktadır:

$$O = \sum_{j \in O} w_j \, s_j \,, \quad E = \sum_{j \in E} w_j \, s_j \tag{3.2}$$

Nihai risk skoru metrik hesaplama motoru tarafından şu şekilde hesaplanmaktadır;

$$R = \frac{O+E}{2} \tag{3.3}$$

Bu çalışmada, standart bir nokta oluşturmak amacıyla eşit ağırlıklandırma stratejisi ($w_O = w_E = 0.5$) kullanılsa da çerçeve, uyarlanabilir ağırlıklandırmayı destekleyecek şekilde tasarlanmıştır. Organizasyonlar, Olasılık ile Etki arasındaki katkıyı kendi risk iştahlarına göre kalibre edebilirler. Örneğin, güvenlik açısından kritik olan bir organizasyon Etki boyutunu daha yüksek bir katsayı ile ifade etmek isterken, tehdit önlemeye odaklanan başka bir organizasyon Olasılığa ağırlık verebilir. Matematiksel olarak çerçeve, $R = \alpha O + (1 - \alpha)E$ genel formunu destekleyecek özelleştirmeye imkan sunmaktadır.

Metrik şemasında tanımlanan operasyonel sınırlar (örn. sektör duyarlılığı, mevzuat etkisi ve temel kurumsal faktörler) kapsamında, her iki boyutta da pozitif alt sınırları koruyabilmektedir. Buna göre, değerlendirilen durumlar için $O>0$ ve $E>0$ ve dolayısıyla $R>0$ sonuçları elde edilmektedir. Okunabilirliği artırmak, anlaşılabilirliği yükseltmek amacıyla skor, $R_{10} = 10R$ formülü ile 0 -10 ölçeğinde raporlanmaktadır.

3.5.3 Alt Parametre Metrik Kalibrasyonu

Her bir alt parametreler için puanlama değerleri; uzmanların 0 - 10 ölçeği üzerinden verdikleri puanlardan oluşan anket tabanlı bir veri seti ($n > 100$) kullanılarak kalibre edilmiştir. Kalibrasyon süreci üç temel adımdan oluşmaktadır:

- **Birleştirme:** Her bir metrik için uzman puanlarının aritmetik ortalaması (μ) hesaplanmıştır.

- **Normalizasyon:** Elde edilen ortalama değerler, çerçevenin hesaplama mantığıyla uyumlu hale getirmek amacıyla [*0.0, 1.0*] aralığına ($W = \mu / 10$) eşlenmiştir.
- **Teorik Sabitleme:** Anket sorularında doğrudan yer almayan sınır durumları, kapsamı desteklemek amacıyla "Sabitleme Noktaları" olarak (Anchor Points) tanımlanmıştır. (örneğin, "Güvenlik Açısından Kritik Etki" değerinin 10 olarak sunulması).

Tablo 3.1'de sunulan kalibre edilmiş ağırlıklar, uzman grubunun risk faktörlerini önceliklerini özetlemektedir. Örneğin, Etki boyutu içerisinde "Gizlilik Etkisi" ($W_F = 0,2586$) en yüksek ağırlık değerini almış ve Teknik Etkiyi ($W_F = 0,0321$) yaklaşık 8 kat aşmıştır. Benzer şekilde, Tehdit Aktörü kategorisi altındaki Kaynaklar ($W_F = 0,1867$), Motivasyon ($W_F = 0,0771$) parametresine ait ağırlık değerinden daha yüksek hesaplanmıştır. Bu durum, uzman grubunun risk puanlaması yaparken saldırganın niyetinden ziyade kabiliyetine öncelik verdiğini göstermektedir.

3.6 OPERASYONEL SINIRLAR VE "ÖNCE GÜVENLİK" MODELLEME İLKELERİ

Geliştirilen çerçevenin temel tasarım aksiyomu; operasyonel ortamlarda aktif bir varlığın artık riskinin (residual risk) pozitif ($R>0$) olarak modellenmesidir. Bu durum saf bir teorik varsayımdan ziyade, çerçevenin "Önce Güvenlik" (Safety-First) mimarisiyle uyumu için bilinçli bir tasarım tercihidir. Mevcut koşullarda, bir risk puanının sıfır (*0,00*) olarak atanması; "sıfır bakım" veya "sıfır tehdit" durumuna karşılık gelecektir. Sürekli izleme gereksinimleri ve bilinmeyen zafiyetlerin (sıfırıncı gün) olasılığı göz önüne alındığında, bu durumun gerekçelendirilmesi oldukça zorlaşmaktadır.

$S_{min}(k)$, k parametresi için mümkün olan en düşük şiddet puanını temsil etmektedir. Uzman görüşüne dayalı şemada tanımlandığı üzere (Bkz: Tablo 3.1 ve Ek A), seçilen bağlamsal parametreler artık riski nicelleştirmek amacıyla sıfırdan farklı bir operasyonel alt sınır getirmektedir.

Etki (*E*) boyutu içerisindeki "Sektör Duyarlılığı" (sens) parametresi, sektörün kritikliğini yansıtan bir temel puan tanımlamaktadır:

$$E_{min} \geq w_{sens} \cdot S_{min}(Sens) = 0{,}1753 \times 0{,}20 = 0{,}035 \quad (3.4)$$

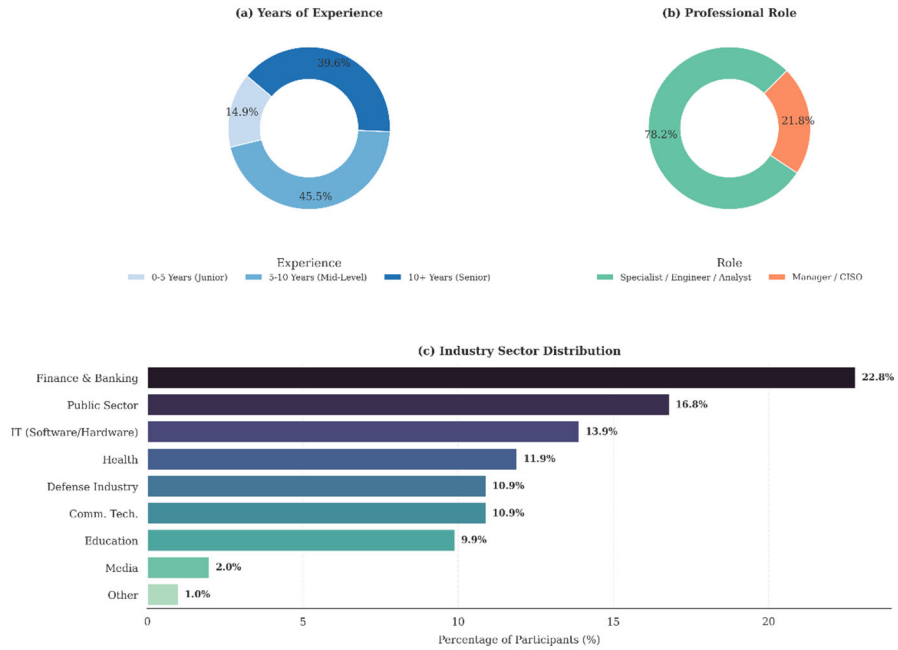
Benzer şekilde, Olasılık (*O*) boyutu için "Yama Mevcudiyeti" gibi teknik parametreler, "Yama Tamamen Uygulanmış" (patch) durumunda bile *0,25* oranında bir artık puanı korumaktadır. Çerçeve, "Tamamen Yamalanmış" durumunu işlevsel bir risk eşdeğeri olarak ele alınmaktadır. Söz konusu bu durum bir zafiyetin; yazılım iyileştirmesi, fiziksel izolasyon ya da sanal yamalama gibi aynı telafi edici kontrollerle etkili bir şekilde zafiyetler yönetilse bile temel artık riski temsil etmektedir. Böyle bir tasarım tercihi, doğrulama maliyetini ve riskin azaltıldığı durumlarda bile atlatılma (bypass) potansiyelini modellemektedir:

$$O_{min} \geq w_{patch} \cdot S_{min}(Patch) = 0{,}1078 \times 0{,}25 = 0{,}027 \quad (3.5)$$

Buna bağlı olarak, Risk Skoru (*R*) için alt sınır pozitif kalmakta ve minimum bir ihtiyat eşiği oluşturmaktadır:

$$\lim_{mitigation \to \infty} R = \frac{O_{artık} + I_{artık}}{2} > 0 \quad (3.6)$$

Bu operasyonel taban, modelin riski yalnızca tek bir değişkene dayanarak (örneğin, bir tarama sonucunu "Hatalı Pozitif" (False Positive) olarak sınıflandırarak) sıfıra indirgenmesini engellemeye yardımcı olmaktadır. Bunun yerine, nihai puanın "Derinlemesine Savunma" (Defense in Depth) ilkesiyle tutarlı bir şekilde sektör bağlamı ve varlık maruziyetiyle ilişkili temel riski yansıtmasını sağlamaktadır.



Şekil 3.2 Anket katılımcılarının demografik profili

3.7 KATILIMCI DEMOGRAFİSİ VE UZMANLIK PROFİLİ

Türetilen risk ağırlıklarının güvenilirliğini desteklemek amacıyla anket, siber güvenlik profesyonellerinden oluşan bir gruba (*N=101*) uygulanmıştır. Şekil 3.2'de ayrıntılarıyla sunulduğu üzere, katılımcı dağılımı çalışma kapsamındaki uzmanlar odağındaki örneklemini yansıtmaktadır:

- **Kıdem:** Şekil 3.2(a)'da görüldüğü üzere, katılımcıların %85,1'i beş yıldan fazla deneyime sahip olduğunu, yaklaşık %40'ı ise uzman düzeyinde (10+ yıl) kıdeme sahip olduğunu bildirmiştir.
- **Karar Verme Yetkisi:** Şekil 3.2(b), katılımcıların %78,2'sinin teknik uzman (mühendis/analist), %21,8'inin ise yönetici veya CISO rollerinde olduğunu göstermektedir. Bu durum, ağırlıkların hem taktiksel hem de stratejik perspektifleri kapsadığını ortaya koymaktadır.
- **Sektörel Çeşitlilik:** Şekil 3.2(c), yüksek düzeyde denetlenen (regulated) endüstrilerin temsil edildiğini göstermektedir; bu

kapsamda Finans ve Bankacılık (%22,8) ile Kamu Sektörü (%16,8) en büyük grupları oluşturmaktadır.

3.8 ETİK HUSUSLAR VE AYDINLATILMIŞ ONAM

Anket, tamamen gönüllülük esasına dayalı olarak gerçekleştirilmiştir. Veri toplama sürecinden önce tüm katılımcılar çalışmanın amacı hakkında bilgilendirilmiş ve onayları alınmıştır. Süreç boyunca hiçbir kişisel bilgi (PII, Kişisel Olarak Tanımlanabilir Bilgiler) toplanmamış, tüm yanıtlar anonim olarak kaydedilmiş ve analiz edilmiştir.

Anket, katılımcının siber güvenlik risk değerlendirmesine ilişkin profesyonel uzmanlık ve teknik tecrübesine odaklanmıştır. Toplanan bilgilerin hassas olmayan veriler içermesi ve insanlar üzerinde deneysel bir çalışma yürütülmemesi nedeniyle, herhangi etik kurul onayı alınmasına gerek duyulmamıştır.

BÖLÜM 4

4. ÖNERİLEN ÇERÇEVE MİMARİSİ

Önerilen çerçeve, dinamik siber risk değerlendirmesi için modüler ve katmanlı bir mimari sunmaktadır. Manuel girdilere dayanan statik metodolojilerin aksine bu mimari, anket tabanlı bir metrik şeması ile sınırlandırılmış, halka açık LLM ajanları (GPT-4o ve Gemini 2.0 Flash) kullanarak otomatik veri alımı, muhakeme ve puanlamayı desteklemektedir.

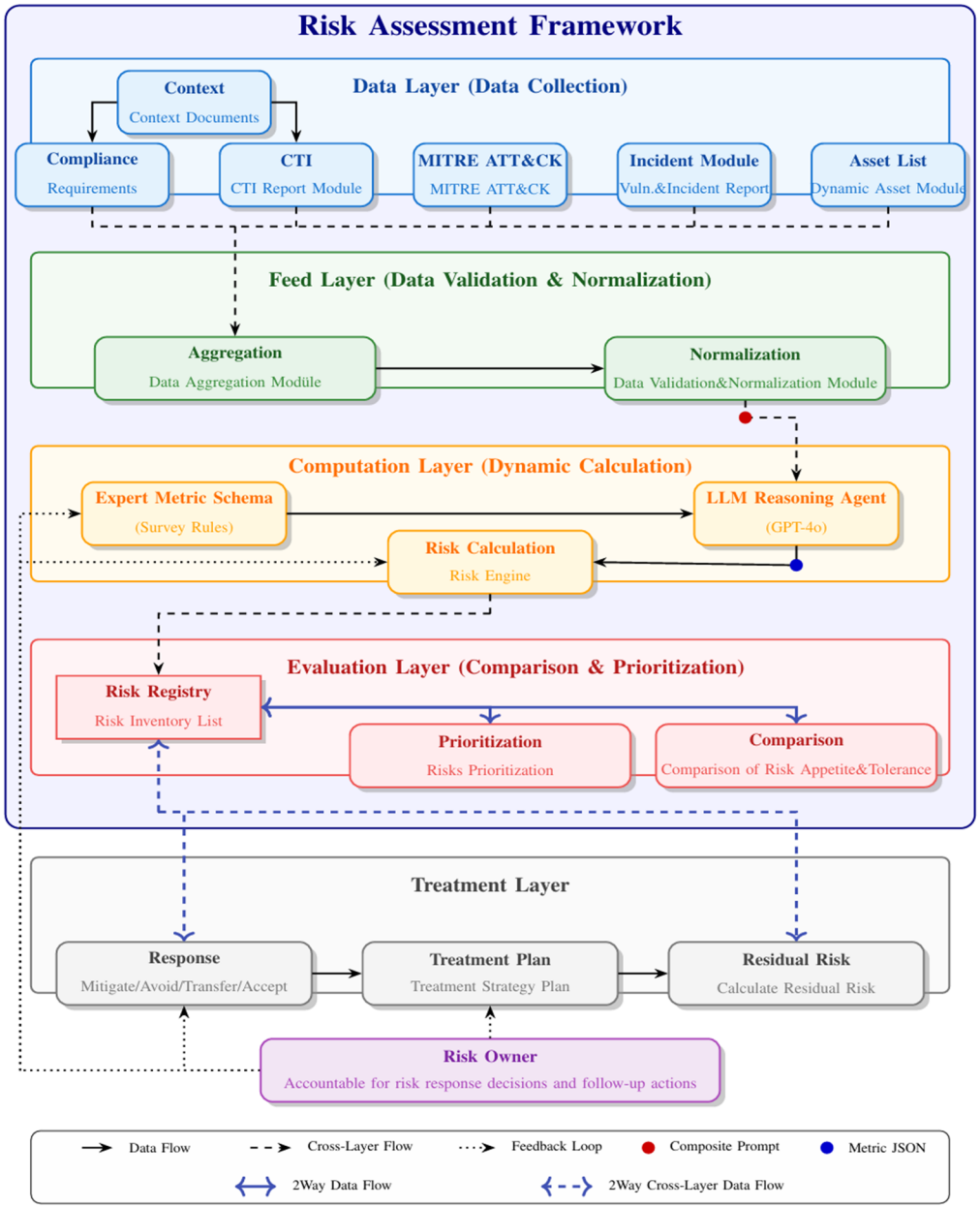
4.1 Genel Mimari

Şekil 4.1'de gösterildiği üzere, önerilen çerçeve ISO 31000 Risk Yönetimi Kılavuzu ile uyumlu olup, veri alımından risk işlemeye kadar yapılandırılmış bir iş akışını tanımlamaktadır. Mimari, birlikte çalışabilir beş katmandan oluşmaktadır:

1. **Veri Katmanı (Tanımlama ve Toplama):** Bu temel katman, kapsamlı bir risk bağlamı oluşturmak amacıyla farklı kaynaklardan gelen heterojen verileri toplamak üzere tasarlanmıştır. Mimari; kurumsal belgelerin, uyumluluk gereksinimlerinin, CTI raporlarının ve sınıflandırma şemalarının alımını desteklemektedir. Fakat bu doğrulama çalışması kapsamında veri toplama süreci, zafiyet raporları ve varlık profilleri ile bilinçli olarak sınırlandırılmıştır. Bu kasıtlı kapsam belirleme, deneysel kontrolü önceliklendirmeye ve yapay zeka muhakeme aşamasındaki değişken gürültüsünü minimize ederek tekrarlanabilirliği artırmaya hizmet etmektedir.
2. **Besleme Katmanı (Doğrulama ve Normalizasyon):** Bir tür ön işleme motoru görevi gören bu katman, ham girdileri yapay zeka ajanının bağlam penceresine (context window) uygun olarak,

metin tabanlı ve yapılandırılmış bir formata dönüştürmektedir. İş akışının şemasında detaylandırıldığı üzere (Bkz: Şekil 4.2, Aşama 1), bu süreç dağınık verileri normalize edilmiş istem bağlamlarına (User Prompt, örn. *vuln.txt* ve *org_context.txt*) dönüştürmektedir. Bu yaklaşım, temiz ve sade metin tabanlı bağlam enjeksiyonu (prompt) elde edilmesine olanak sağlar. Yöntem token (LLM sorgularının üçretlendirilmesi) maliyetini azaltırken, yapılandırılmış JSON formatları ise LLM yanıtlarının kesin ve istenen yapıda üretilmesine hizmet eder.

3. **Hesaplama Katmanı (Risk Hesaplaması):** Çerçevenin merkezinde yer alan LLM muhakeme ajanı, hazırlanan bağlamı (prompt) analiz etmekte, uygun alt parametreleri seçmekte ve bir JSON çıktısı üretmektedir. Nihai risk skoru ise, Tablo 3.1'de tanımlanan uzman görüşlü ağırlıklar kullanılarak hesaplanmaktadır. Bu katman aynı zamanda bir HITL (Human-in-the-Loop / Süreçte İnsan) mekanizması içermektedir. Risk Sahibi, geri bildirim sağlayabilmekte ve gerektiğinde metrik seçimlerini manuel olarak ayarlayabilmektedir.
4. **Değerlendirme Katmanı (Karşılaştırma ve Önceliklendirme):** Hesaplama süreci tamamlandıktan sonra tanımlanan riskler merkezi Risk Kayıt Defteri'ne (Risk Registry) kaydedilmektedir. Sistem, bu riskleri hesaplanan şiddet puanlarına ve kurumun tanımlanmış risk iştahına göre önceliklendirerek yapılandırılmış iyileştirme planlamasını desteklemektedir.
5. **Risk İşleme Katmanı (Yanıt):** Son aşamada çerçeve, önceliklendirilmiş riskleri Risk Sahibi' ne sunmaktadır. Risk Sahibi uygun müdahale stratejisini (Azalt, Transfer Et, Kabul Et veya Kaçın) seçmektedir. Bu karar, iyileştirme akışlarını tetiklemektedir.



Şekil 4.1 Risk Değerlendirme Çerçevesi Genel Mimari

4.2 Veri Katmanı: Tanımlama ve Toplama

Veri Katmanı, çerçevenin temelini oluşturmaktadır. Bu katman, risk değerlendirmesiyle ilgili yapılandırılmış ve yapılandırılmamış bilgi kaynaklarını derlemektedir:

- **Bağlamsal Bilgiler:** Kurumsal politikalar, Hizmet Seviyesi Anlaşmaları (SLA) ve iş açısından kritiklik faktörleri.
- **CTI Beslemeleri:** Tehdit göstergeleri ve saldırgan profilleri sunan siber tehdit istihbaratı raporları (örn. MISP, OpenCTI).
- **MITRE ATT&CK:** Davranışsal tehdit modellemesini bilgilendiren taktikler, teknikler ve prosedürler (TTP'ler) (Al-Sada vd., 2024).
- **Olay Raporları ve Varlık Envanteri:** Varlık kritikliğini, maruziyeti ve aktif güvenlik olaylarını açıklayan dinamik kayıtlar.

4.3 Besleme Katmanı: Doğrulama ve Normalizasyon

Besleme Katmanı, ham girdiler ile YZ motoru arasında tampon görevi görmektedir. Katmanın temel işlevi ön işlemedir. Veriler büyük dil modeline (LLM) sunulmadan önce ham metinler temizlenmekte ve heterojen kayıtları birleşik bir JSON yapısına dönüştürülmektedir. Süreç LLM' nin gürültülü ham veriler yerine yapılandırılmış bağlam enjeksiyonu (context injection) kullanmasına imkan tanımakta, halüsinasyon (hallucination) riskini en aza indirgemektedir.

4.4 Hesaplama Katmanı: LLM Tabanlı Akıl Yürütme

Çerçevenin merkezinde, geleneksel manuel analizini yapay zeka muhakeme ajanıyla ikame eden Hesaplama Katmanı yer almaktadır. Uçtan uca, kapalı kutu (black-box) derin öğrenme modellerine dayanmak yerine farklı bir iş akışı tercih edilmiştir. Bu kapsamda, modelin sabit ve önceden işlenmiş bir bağlam paketi aldığı, mantıksal olarak RAG (*Retrieval-Augmented Generation*) sistemlerine benzer bir bağlam enjeksiyonu (*context-injection*) yöntemi kullanılmaktadır. Özellikle, değerlendirme sürecinde harici bir bilgi getirme veya farklı araçlar ihtiyaç duyulmamıştır.

Çalışmadaki temel bir ayrım, Algoritmik Opaklık (Algorithmic Opacity) ile Sistemik Şeffaflık (Systemic Transparency) arasındadır. Temeldeki LLM' in (GPT-4o veya Gemini 2.0 Flash) metirkleri seçimi açısından ticari bir "kapalı kutu" olsa da; LLM' in metrikleri önceden tanımlanmış kategorilere eşlemek için gerçekleştireceği anlamsal çıkarım (semantic extraction) sınırlamaktadır. Ayrıca riskin hesaplanması, Metrik Motoru tarafından yönetilen şeffaf ve deterministik yapı ile korumaktadır. Mimari, anlamsal işleme bileşenini (metrik seçimi) puanlama mantığından (hesaplama) ayırmaktadır.

YZ ajanı, ham girdileri nicel metricler ile eşleyen yapılandırılmış, dört aşamalı bir iş akışını (pipeline) izlmektedir (Bkz: Şekil 4.2).

4.4.1 Aşama 1: Ön İşleme (Besleme Katmanı)

Ham girdiler; belirteç (token) gürültüsünü azaltmak ve LLM yorumlamasında tutarlılığı sağlamak amacıyla ilk olarak normalize edilmektedir. Bu aşamada, zafiyet ilişkin heterojen bilgiler standartlaştırılmış bir "*vuln.txt*" dosyasına dönüştürülürken (Bkz: Ek B, Şekil B.1), kurumsal varlıklar ve organizasyon bağlamına ilişkin bilgiler "*org_context.txt*" dosyasında özetlenmektedir (Bkz: Ek B, Şekil B.2). Bu süreç, YZ ajanına ham veri akışları yerine yapılandırılmış metin tabanlı istem bağlamları (prompt contexts) sağlanarak, anlamsal çıkarım için hedeflenmiş stratejiyi işaret etmektedir. Tekrarlanabilirliği desteklemek ve sıkı deneysel kontrolü sürdürmek amacıyla bu çalışmadaki ön işlem kapsamı kasten temel bilgiler ile sınırlandırılmıştır. Bu şekilde sonraki YZ muhakeme aşaması için daha güvenilir bir temel oluşturulmasını sağlamaktadır.

4.4.2 Aşama 2: Bağlam Enjeksiyonu

Karma bir istem (composite prompt) yapısı için LLM' e dört ana bileşen gönderilir:

- **Sistem Kimliği:** YZ' nin "Kıdemli Siber Risk Analisti" olarak tanımlanmasıdır.

- **Metrik Şeması:** "Metrik Kutusu"nun (Bkz: Tablo 3.1) JSON formatıdır. Parametre ve metriklere ait tanımları içerir.
- **Bağlam Verisi:** Aşamada 1' de hazırlanan, önceden işlenmiş metin dosyalarıdır.
- **Mantıksal Kurallar (Logic Overrides):** "Önce Güvenlik" (safety-first) yaklaşımını desteklemek amacıyla; belirsiz sınır durumlarını (örneğin, fiziksel olarak izole edilmiş/air-gapped varlıklar) yönetmeye yönelik bir dizi deterministik kural kümesidir (Bkz: Ek A, Şekil A.2).

4.4.3 Aşama 3: YZ Akıl Yürütme ve İhtiyatlı Protokol

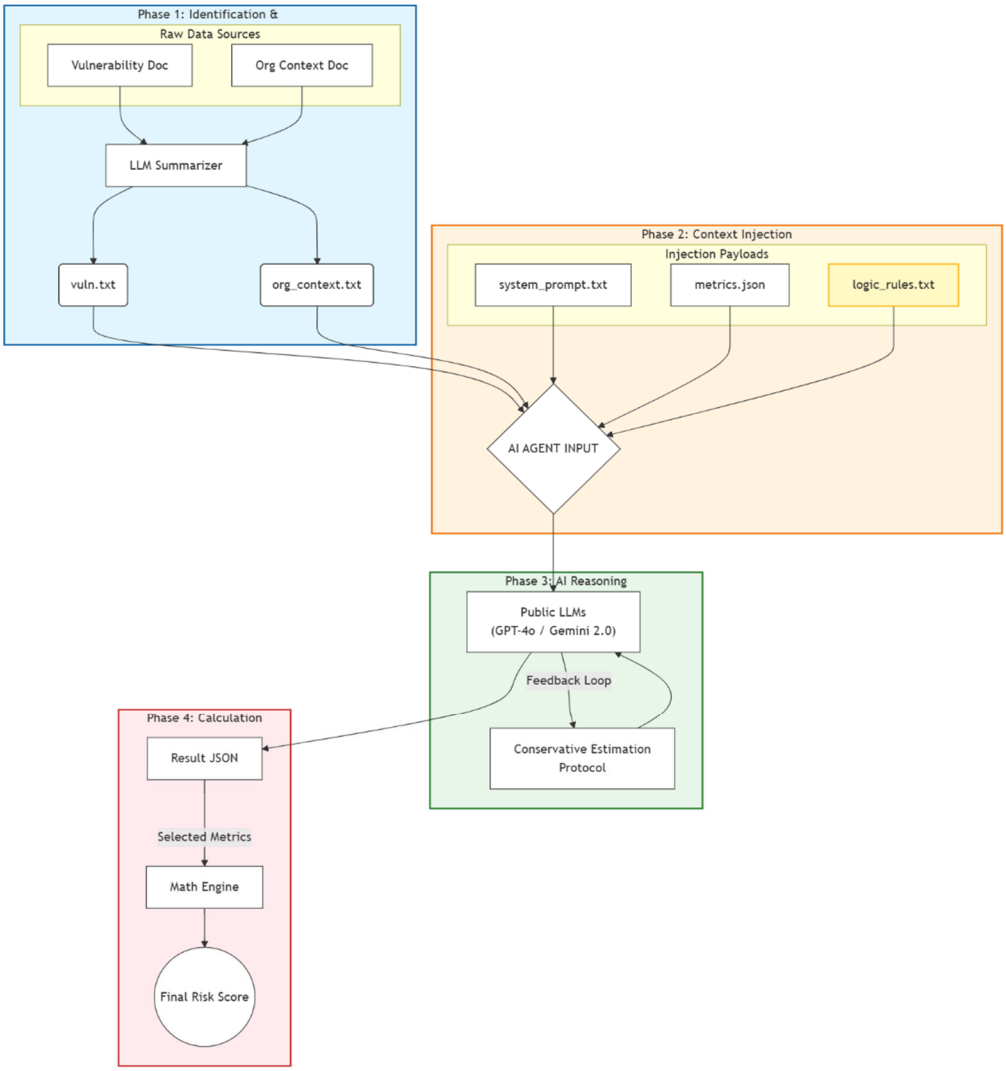
LLM (GPT-4o veya Gemini 2.0 Flash), gönderilen bağlamı analiz etmekte ve parametreleri alt parametrelerle eşlemektedir. Ajan, hatalı negatif (false negative) riskini azaltmak amacıyla, belirsizlik durumlarında daha risk skorunda en yüksek şiddete sahip seçeneği seçmek üzere sınırlandırılmıştır. Bir "İhtiyatlı Tahmin Protokolü" takip etmektedir (Bkz: Şekil 4.2).

4.4.4 Aşama 4: Çıktı ve Hesaplama

LLM, yalnızca seçilen metrik anahtarlarını içeren yapılandırılmış bir JSON dosyası ("*Result JSON*") üretmektedir. Nihai skor hesaplaması, "Matematik Motoru" tarafından ROC ağırlıkları (Bkz: Eşitlik 3.1) kullanılarak harici olarak gerçekleştirilmektedir. Bu yaklaşım sayesinde, LLM tarafından aritmetik işlemler gerçekleştirilmemekte ve olası sayısal hatalar minimize edilmektedir.

4.5 Değerlendirme Katmanı: Puanlama ve Önceliklendirme

YZ ajanı uygun alt parametreleri (örneğin; Saldırı Vektörü: Ağ, Etki: Yüksek) seçtikten sonra, Değerlendirme Katmanı uzman görüşüne dayalı ROC ağırlıklarını (Bkz: Eşitlik 3.1) kullanarak nihai risk skorunu hesaplamaktadır.



Şekil 4.2 YZ Risk Değerlendirme İş Akışı

4.6 Müdahale Katmanı: Yanıt ve Yönetişim

Değerlendirme sürecinin ardından İşleme Katmanı, risk yanıt seçeneklerini (Azalt, Transfer Et, Kabul Et) sunarak risk yönetimini destekler. LLM, tanımlanan zafiyete dayalı olarak iyileştirme eylemleri önerebilse de; nihai karar HITL (Human-in-the-Loop / Süreçte İnsan) yönetişimi altındadır.

BÖLÜM 5

5. KARŞILAŞTIRMALI DOĞRULAMA: İNSAN UZMANLAR VE YZ AJANI PERFORMANSI

Önerilen çerçevenin güvenirliliğini ve ölçeklenebilirliğini değerlendirmek amacıyla karşılaştırmalı çalışma yürütülmüştür. Yalnızca sentetik verilere dayanan önceki çalışmaların aksine bu çalışmada gerçek dünya zafiyet senaryolarını kullanılarak bir grup kıdemli uzman ile YZ ajanı karşılaştırmalı olarak değerlendirilmiştir. Çalışmanın temel amacı çerçevenin puanlama doğruluğu, tutarlılık ve operasyonel verimlilik olmak üzere üç ana boyutta incelenmesidir.

5.1 Deney Tasarımı ve Senaryo Bazlı Veri Kümesi

Tablo 5.1'de ayrıntılandırıldığı üzere bağlam uyarlanabilirliğini değerlendirmek üzere bir temel oluşturmak amacıyla, Kritik Altyapılardan (C_8, C_{15}) KOBİ'lere (C_{13}) kadar geniş yelpazede on beş senaryodan (C_1-C_{15}) oluşan bir veri seti kullanılmıştır. Veri seti ayrıca iki doğrulama akışına ayrılmıştır:

- **Tarihsel Olayların Değerlendirilmesi (C_1-C_{15}):** Finans, Savunma, Enerji ve Sağlık dahil olmak üzere kritik sektörleri kapsayan, yaygın olarak belgelenmiş ve gerçek dünyada yaşanmış siber güvenlik olaylarından oluşan bir settir. Bu akış, model çıktılarını tarihsel olarak raporlanan sonuçlarla karşılaştırmak amacıyla yüksek etkili fidye yazılımı olaylarını (örneğin; Colonial Pipeline, C_8) karmaşık tedarik zinciri saldırıları ve mantıksal kusurlarla (örneğin; SolarWinds, Optus) birleştirmektedir.
- **Karşılaştırmalı Bağlam Analizi (C_{13}-C_{15}):** Gerçek dünyadaki teknik bir zafiyetin (yamalanmamış RCE), üç farklı kurumsal

bağlamda (KOBİ, Büyük Ölçekli İşletme ve Kritik Altyapı) uygulandığı kontrollü bir karşılaştırmalı çalışmadır. Bu akış, çerçevenin aynı teknik kusur için risk puanlarını nasıl uyarladığını incelemek amacıyla kurumsal bağlamın etkisini ayrıştırmaktadır.

Tablo 5.1 Gerçek Dünya Vaka Çalışmaları (C1--C15)

Vaka No	Sektör ve Bağlam	Zafiyet Profili	Varlık Kritikliği
C_1	**Finans (Merkez Bankası):** Ulusal Kritik Altyapı	Ağ Segmentasyonu Hatası, Günlük Kaydı Eksikliği ve APT Zararlı Yazılımı (Mandiant, 2023)	Kritik (Egemen Fonlar ve SWIFT Kimlik Bilgileri)
C_2	**Savunma Sanayii:** Stratejik Ulusal Güvenlik Kuruluşu	Komut Enjeksiyonu (*Prompt Injection* - OWASP LLM01) ve Bozuk Erişim Kontrolü	Kritik (Gizli Tasarım Şemaları ve Ar-Ge Verileri)
C_3	**Havacılık / Savunma:** Küresel 1. Seviye Savunma Yüklenicisi (The Record, 2023b)	Yamalanmamış Zafiyet (Citrix Bleed) ve Fidye Yazılımı (LockBit 3,0) (The Record, 2023a)	Kritik (Tasarım Fikri Mülkiyeti, Gizli Veriler ve Filo Desteği)
C_4	**Sağlık / Kritik Altyapı:** Sağlık Hizmetleri için SaaS Sağlayıcısı	Fidye Yazılımı, Yamalanmamış Zafiyet (ZeroLogon) ve MFA Eksikliği (CISA, 2021)	Yüksek (Özel Nitelikli Tıbbi Veriler)
C_5	**E-Ticaret (B2C):** Büyük Ölçekli Perakendeci	Uzaktan Kod Çalıştırma (RCE) için XXE (CosmicSting) (Sansec, 2024)	Kritik (Müşteri Kişisel Verileri, Ödeme Tokenizasyonu)
C_6	**Küresel SaaS / Bulut:** Halka Açık Kurumsal Şirket	Bozuk Kimlik Doğrulama (Mantık Hatası) ve Yetki Yükseltme (ATO) (CrowdStrike, 2023)	Kritik (Kimlik Yönetimi ve Finansal Veriler)
C_7	**Siber Güvenlik:** Kitle Kaynaklı Pazaryeri	İş Mantığı Hatası ve IDOR (İtibar Manipülasyonu) (HackerOne, 2023)	Kritik (İtibar Sistemi Bütünlüğü ve Platform Güveni)

Tablo 5.2 (Devamı) Gerçek Dünya Vaka Çalışmaları (C1--C15)

Vaka No	Sektör ve Bağlam	Zafiyet Profili	Varlık Kritikliği
C_8	**Enerji (Kritik Altyapı):** Yakıt Boru Hattı İşletmecisi	Sızdırılmış VPN Kimlik Bilgisi (MFA Yok) ve Fidye Yazılımı (DarkSide) (CISA, 2021; Miller, 2021)	Kritik (Yakıt Tedariki ve Faturalama Sürekliliği)
C_9	**Tedarik Zinciri (Yazılım):** Ulusal Güvenlik Satıcısı	Ele Geçirilmiş Derleme Hattı (Sunburst Arka Kapısı) (CISA, 2020; Microsoft, 2021)	Kritik (Hükümet Ağları ve Kurumsal Güven)
C_{10}	**Konaklama (Turizm):** Küresel Casino ve Otel İşletmecisi	Sosyal Mühendislik (Yardım Masası Üzerinden Sesli Kimlik Avı) ve IdP İhlali (SEC, 2023; CyberArk, 2023)	Yüksek (Misafir Hizmetleri ve Casino Operasyonları)
C_{11}	**Otomotiv (İmalat):** JIT Tedarik Zinciri	Üçüncü Taraf Tedarikçi İhlali (JIT Hatası) (Nikkei, 2022; The Record, 2022)	Kritik (Üretim Hattı Sürekliliği)
C_{12}	**Telekomünikasyon:** Ulusal Operatör	Kimlik Doğrulaması Yapılmamış API Uç Noktası (BOLA/IDOR) (OAIC, 2022)	Kritik (Müşteri Kişisel Verileri - 9,8 Milyon Kayıt)
Bağlam Duyarlılığı Senaryoları (Özdeş Teknik Zafiyet: Log4Shell CVE-2021-44228, CVSS: 10,0)			
C13	**KOBİ (Lojistik):** Mutfak Menüsü Taslak Bilgisayarı	Varlık hava boşluklu (air-gapped) ve fiziksel olarak izole edilmiştir; hassas veri içermez.	Düşük (Yalnızca operasyonel iş yükü)
C14	**Büyük Ölçekli İşletme:** Dahili Kurumsal Ağ	Sunucu, çalışan portallarını ve gizli kişisel verileri içeren entegre ödeme geçitlerini barındırır.	Yüksek (Dahili Operasyonlar ve Gizlilik)
C15	**Kritik Altyapı:** Yaşam Destek IoT Geçidi	Ventilatörlerden gelen telemetriyi toplar; veri kaybı doğrudan hasta güvenliğini etkiler.	Kritik (Kamu Güvenliği ve Ulusal Güvenlik)

5.1.1 YZ Ajanının Yapılandırılması ve Sonuçların Tutarlılığı

Değerlendirmeler, model agnostikliğini incelemek amacıyla, halka açık iki büyük dil modeli (LLM) üzerinde gerçekleştirilmiştir:

(i) GPT-4o (OpenAI) ve

(ii) Gerçek zamanlı işleme yeteneklerini değerlendirmek üzere seçilen Gemini 2.0 Flash (Google).

Her iki modelde de kararlı (düşük varyanslı) çıktılar elde etmek amacıyla *T=0,1* sıcaklık değeri ile yapılandırılmıştır, Bölüm 4.4'de açıklanan metin tabanlı bağlam enjeksiyonu (context-injection) iş akışı kullanılmıştır. Tekrarlanabilirliği desteklemek ve sıkı deneysel kontrolü sürdürmek için girdi olarak kullanılan veriler, kasıtlı olarak zafiyet raporları ve varlık profilleriyle sınırlandırılmıştır. Sonuçlar, on kişiden oluşan kıdemli uzman kontrol grubu (N=10) ile karşılaştırılmıştır.

LLM tabanlı tüm değerlendirmeler, sabit değerlendirme penceresi (24.11.2025 ile 29.11.2025 arası) içerisinde yürütülmüştür. Model güncellemeleri karşısında izlenebilirliği sağlamak amacıyla çalışma sırasında her bir sağlayıcı için model kimliği/sürüm dizisi (örneğin; *gpt-4o-<versiyon>* ve *gemini-2.0-flash-<versiyon>*) ve çalışma ayı (*2025-11*) kaydedilmiştir. Kod çözme (decoding) parametreleri tüm çalışmalarda sabit tutulmuştur:

- *temperature = 0,1*
- *top_p = 1,0*
- *max_tokens = 512*
- *presence_penalty = 0*
- *frequency_penalty = 0*

Herhangi bir araç/fonksiyon çağırma özelliği etkinleştirilmemiş ve Bölüm 4.4'de açıklanan sabit metin tabanlı bağlam enjeksiyonu dışında harici bir bilgi getirme (retrieval) işlemi kullanılmamıştır.

Tüm vakalar için sistem istemi (prompt), metrik şeması (Tablo 3.1) ve ihtiyatlı seçim protokolü sabitlenmiştir. Tam istem şablonu ve JSON çıktı şeması makul bir talep üzerine yazarlardan temin edilebilmektedir. Tüm deneylerde

değişmez sistem talimatları (Bkz: Ek A) ve standartlaştırılmış girdi bağlamları (Bkz: Ek B) kullanılmıştır. Bu durum, olası model güncellemelerine rağmen karşılaştırılabilirliği korumayı amaçlayan bir metodolojiye işaret etmektedir.

5.1.2 Uzman Değerlendirme Temeli ve Ölçüm

On kıdemli siber güvenlik uzmanından (CISSP/CISM sertifikalı, ortalama 12 yıllık deneyim) (Bkz. Tablo 5.2) oluşan bir kontrol grubu, aynı 15 senaryoyu değerlendirmiştir. Bu değerlendirmenin amacı, siber risk için mutlak bir "temel gerçeklik" iddiasında bulunmak değildir. Önerilen metrik şemasının tutarlı uzman görüşünü destekleyecek yeterli açıklığı sağlayıp sağlamadığını ve LLM ajanlarının bu şemayı ölçeklenebilir şekilde operasyonelleştirip operasyonelleştiremeyeceğini test etmektir.

Tablo 5.1 Doğrulama Çalışmasına Katılan Uzmanların Profili

Uzman No	Rol / Ünvan	Deneyim (Yıl)
E1	CISO / Bilgi Güvenliği Müdürü	18
E2	CISO / Bilgi Güvenliği Müdürü	15
E3	Risk Yönetimi / Uyum Uzmanı	12
E4	CISO / Bilgi Güvenliği Müdürü	18
E5	Uygulama Güvenliği Uzmanı	9
E6	SOC / Olay Müdahale Analisti	7
E7	Risk Yönetimi / Uyum Uzmanı	10
E8	SOC / Olay Müdahale Analisti	9
E9	Uygulama Güvenliği Uzmanı	8
E10	Risk Yönetimi / Uyum Uzmanı	12

Algoritmik karşılaştırma için bir temel oluşturmak amacıyla, uzmanlardan sezgisel risk puanları vermeleri istenmemiştir. Bunun yerine uzmanlara, önerilen "Metrik Kutusu" da tanımlı parametreler için uygun metriklerin seçimlerini yapmaları talimatı verilmiştir (Bkz: Tablo 3.1). Bu kontrollü protokol, standartlaştırılmış bir analitik yapı sunarak öznel değişkenliği minimize etmeyi amaçlamıştır. Yapılan analizler sonucunda, 0,996 gibi oldukça yüksek bir

değerlendiriciler arası güvenirlik (Inter-rater Reliability) puanı (Cronbach α=0,996) elde edilmiştir. Bu sonuç, geliştirilen şemanın karmaşık risk parametreleri genelinde uzman yargısını disipline edebilecek düzeyde açık ve anlaşılır olduğunu ortaya koymaktadır.

5.1.3 Bağlamsal Değişkenlerin Etkisini Ölçen Kontrollü Deney (C_{13} - C_{15})

Tarihsel olayların geriye dönük analizleri olan C_1-C_{12} senaryolarının aksine, C_{13}-C_{15} senaryoları "Kurumsal Bağlam" etkisini ayrıştırmak amacıyla kontrollü değişken testi olarak kurgulanmıştır. Buradaki temel amaç, önerilen çerçevenin sabit kalan "teknik şiddet" ile bağlama göre değişen "iş riski" arasındaki ayrımın yapılıp yapılamadığını incelemektir.

Bu üç vaka için YZ ajanına sağlanan teknik zafiyet girdisi sabitlenmiş ve özdeş tutulmuştur:

"SRV-WEB-01 adlı dışa açık web sunucusunda, kimlik doğrulaması gerektirmeyen kritik bir Uzaktan Kod Çalıştırma (RCE) zafiyeti (CVSS Temel Puanı: 9,8) tespit edilmiştir. İstismar kodu (exploit) kamuya açıktır."

(Orjinal metin: "A critical unauthenticated Remote Code Execution (RCE) vulnerability (CVSS Base Score: 9.8) has been detected on the external-facing web server 'SRV-WEB-01'. The exploit is publicly available.")

Buna karşılık gönderilen kurumsal bağlam belgeleri şu şekilde çeşitlendirilmiştir:

- **C_{13} (KOBİ Bağlamı):** Varlık, yerel bir işletme için statik bir tanıtım web sitesi olarak tanımlanmıştır. Servis bağımsız bir VPS üzerinde barındırılmaktadır. Hassas veri içermemekte ve sunucunun iç ağlara yönelik yatay hareket (*lateral movement*) yolları bulunmamaktadır.
- **C_{14} (Büyük Ölçekli İşletme Bağlamı):** Varlık, büyük bir işletmenin birincil e-ticaret portalı olarak tanımlanmıştır. Arka planda uç ödeme geçitleriyle entegre durumdadır ve müşteri PII verilerini içermektedir.
- **C_{15} (Kritik Altyapı Bağlamı):** Varlık, bölgesel bir elektrik dağıtım birimi için bir BT/OT geçidi (*gateway*) olarak işaretlenmiştir. Bu sunucunun ele geçirilmesi, saldırganların SCADA ağına sızmasına ve

potansiyel olarak kamu hizmetlerinde fiziksel kesintiye neden olmasına yol açabilmektedir.

Bu tasarım bağlam değişkenini ayrıştırmakta, bu sayede nihai risk skorunda gözlemlenen tüm varyasyonlar, ajanın varlık kritikliğini ve etki bağlamını yorumlama yeteneğini doğrudan gösterebilmektedir.

5.2 Ardışık Testlerde Üretilen Sonuçların Kararlılığı

Çıktı kararlılığını değerlendirmek amacıyla her bir senaryo; özdeş girdiler (aynı bağlam dosyaları, sistem istemi ve metrik şeması) ve özdeş kod çözme (decoding) parametreleri (*temperature=0,1 , top_p =1,0*) altında, her bir LLM model için N=10 kez yeniden çalıştırılmıştır. Bu kapsamda;

(i) seçim uyumu (selection agreement) ve

(ii) nihai skorun standart sapması (standard deviation) ile ölçülen skor değişkenliği (score variability) raporlanmıştır.

Toplu raporlama için, tüm senaryolar (C_1--C_{15}) genelinde ortalaması alınan seçim uyumu hesaplanmıştır. Ayrıca Tablo 5.3'te özetlendiği üzere, senaryolar genelindeki ortalama skor standart sapması sunulmuştur.

Tablo 5.2 Tekrarlanan Çalışmalar Üzerinden Kararlılık Sonuçları (C1--C15)

Model	**Senaryo Başına Çalıştırma Sayısı (N)**	**Seçim Tutarlılığı (%)**	**Puan Standart Sapması**
GPT-4o	10	97,4	0,089
Gemini 2.0 Flash	10	98,2	0,066

5.3 Karşılaştırmalı Analizden Elde Edilen Temel Bulgular

5.3.1 YZ ve İnsan Kararlarındaki Puanlama Uyumu

Şekil 5.1'de görselleştirilen karşılaştırmalı analizde LLM destekli değerlendirmelerin, özellikle yüksek şiddetli senaryolarda insan medyan

değerleriyle yakından uyumlu olduğunu göstermektedir. Doğrulama seti (C_1--C_{15}) genelinde hem GPT-4o hem de Gemini 2.0 Flash, on kıdemli uzmandan oluşan kontrol grubuyla güçlü bir pozitif korelasyon sergilemiştir.

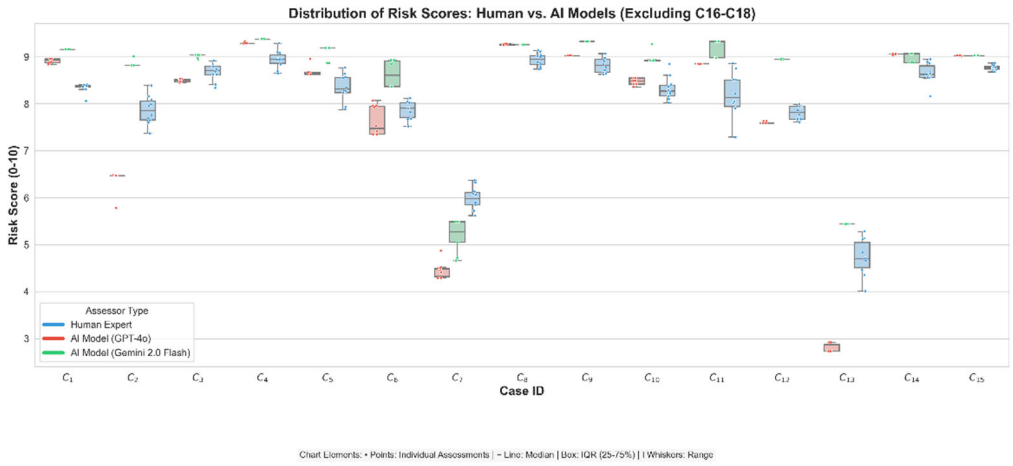
Kritik altyapı senaryolarında, özellikle Sağlık Sektörü Fidye Yazılımı (C_4), Enerji Sektörü Boru Hattı (C_8) ve SolarWinds Tedarik Zinciri İhlali (C_9) vakalarında, YZ ajanları 9,0' ın üzerinde medyan skorlar üretmiştir. Üretilen sonuçların, uzman değerlendirmelerinin merkezdeki %50'lik dilimini temsil eden çeyrekler arası aralık (*interquartile range*) içinde yer alması; modelin, kıdemli analistlerin ortak yargısıyla istatistiksel olarak uyumlu olduğunu göstermektedir.

Elde edilen sonuçlar ayrıca daha düşük bir dağılıma işaret etmektedir. Savunma Sanayii İstem (prompt) Enjeksiyonu (C_2) ve JIT Tedarik Zinciri Hatası (C_{11}) gibi karmaşık vakalarda, insan değerlendiriciler daha geniş bir puan yayılımı sergilemiştir. Bu durum kutu grafiklerinde de görülebilmekte ve yorumlama farklılıklarına işaret etmektedir. Buna karşın YZ ajanları, tekrarlanan çalışmalar genelinde daha dar bir dağılım göstermiştir. Bu örüntü, aynı şema altındaki manuel değerlendirmeye kıyasla öznelliğin azaldığını ortaya koymaktadır. Ayrıca, C_{13} senaryosunda C_{14} vakasına kıyasla gözlemlenen puan düşüşü, ilerleyen bölümlerde daha ayrıntılı incelenecek olan bağlam duyarlılığına dair ilk kanıtları sunmaktadır.

5.3.2 Sonuçların İstatistiksel Bağıntısı ve Güvenilirlik Testleri

Kutu grafiği analizi görsel bir uyum kanıtı sunsa da (Bkz. Şekil 5.1), bu ilişki 15 doğrulama senaryosu (N=15) genelindeki istatistiksel testler aracılığıyla daha ileri düzeyde nicelleştirilmiştir.

İlk olarak, bir temel çizgi oluşturmak amacıyla insan kontrol grubunun güvenilirliği değerlendirilmiştir. Uzman grubu, α = 0,996 Cronbach alfası değeriyle oldukça yüksek bir iç tutarlılık sergilediği görülmüştür. Bu yüksek uyum derecesi, Metrik Kutusu içindeki yapılandırılmış tanımların uzman yargısını etkili bir şekilde disipline edebildiğini göstermektedir.



Şekil 5.3 C1--C15 Genelinde Risk Puanlarının Dağılımı

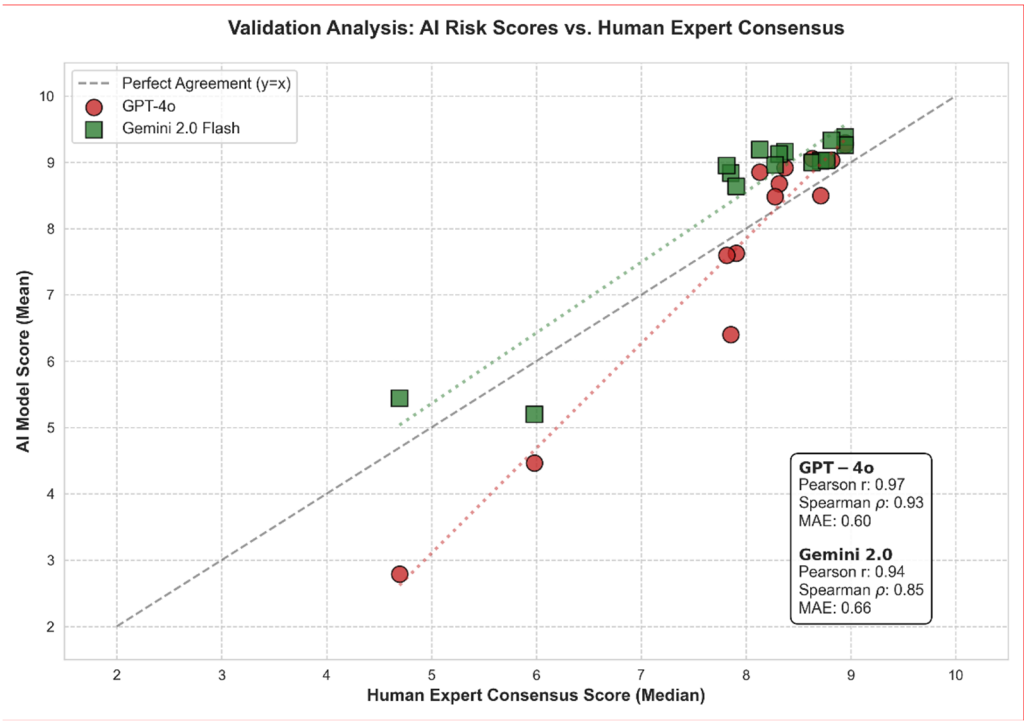
Bu medyan tabanlı temel çizgiye karşı, GPT-4o ajanı güçlü bir doğrusal korelasyon (r = 0,9717 , p < 0,001) ve güçlü bir monotonik ilişki (Spearman ρ = 0,9276) sergilemiştir. Benzer şekilde, Gemini 2.0 Flash ajanı da güçlü bir uyum ortaya koymuştur (r = 0,9390 , p < 0,001 , ρ = 0,8472$). Şekil 5.2'de görselleştirilen bu sonuçlar; Metrik Motorunun, LLM çıktılarını uzman temel çizgisiyle tutarlı risk skorlarına dönüştürebildiğine işaret etmektedir.

Senaryolar arasındaki genel artış/azalış eğilimlerinin ötesinde; YZ ajanları ile insan uzmanların verdiği puanların birbirine sayısal olarak ne kadar yakın olduğu (mutlak uyum) nicel yöntemlerle hesaplanmıştır. 15 doğrulama senaryosu genelindeki Ortalama Mutlak Hata (MAE); GPT-4o için 0,5991, Gemini 2,0 Flash için ise 0,6613 olarak bulunmuştur. 0 ile 10 arasındaki risk ölçeği dikkate alındığında bu değerler, %6,7' den daha az bir ortalama sapmaya tekabül etmektedir. Bu durum, ajanların sadece insan eğilimini takip etmekle kalmayıp, aynı zamanda oldukça düşük hata oranına sahip skorlar ürettiğini göstermektedir.

5.3.3 Tutarlılık ve Kararlılık

LLM tabanlı sistemlerin rastlantısal yapısı, bu sistemlere ilişkin temel bir endişe kaynağı olmaktadır. Ancak, genişletilmiş kararlılık analizi (Bkz: Tablo

5.3), önerilen mimari dahilinde değişkenliğin sınırlı olduğunu göstermektedir. 150 deneysel iterasyon genelinde, GPT-4o %97,4'lük bir Seçim Uyumu (Selection Agreement) sağlarken; Gemini 2.0 Flash %98,2'lik orana ulaştığı görülmüştür. Bu sonuçlar, "Metrik Kutusu"nun halüsinasyona karşı bir koruma bariyeri görevi gördüğünü, ajanların çoğu çalışmada aynı metrik parametrelerini seçtiğini ortaya koymaktadır.

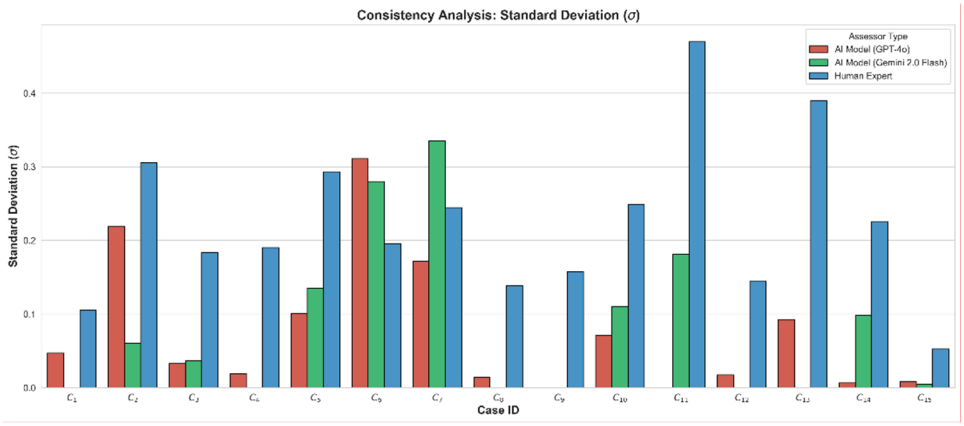


Şekil 5.4 İnsan Medyan-YZ Puanları Arasındaki Korelasyon Analizi

Söz konusu gözlemle tutarlı olarak Şekil 5.3 , YZ ajanlarının düşük skor değişkenliğini koruduğunu ve çoğu vakada bu değerin $\sigma \approx 0{,}15$ altında kaldığını göstermektedir. Bu durum, standart sapmanın sıklıkla $\sigma \approx 0{,}30$ değerini aştığı ve karmaşık senaryolarda $\sigma \approx 0{,}45$ seviyesine ulaştığı uzman temel çizgisine nazaran, önemli ölçüde kararlılığı göstermektedir.

Bu ayrışma, özellikle KOBİ Bağlamı (C_{13}) ve Savunma İstem Enjeksiyonu (C_2) gibi vakalarda belirginlik göstermektedir. Bu senaryolarda insan uzmanlar, nihai şiddet seviyesi konusunda daha büyük bir görüş ayrılığı sergileyerek daha

yüksek varyans sonuçları üretmiştir. Buna karşın hem GPT-4o hem de Gemini 2.0 Flash, tekrarlanan çalışmalarda oldukça düşük sapma görülmüştür. Genel olarak önerilen çerçeve; yüksek uyarı hacmi ve zaman kısıtlı değerlendirme koşulları altında, tekrarlanabilir risk puanlamasını desteklemektedir.



Şekil 5.5 Tutarlılık Analizi: Puanların Standart Sapması (σ)

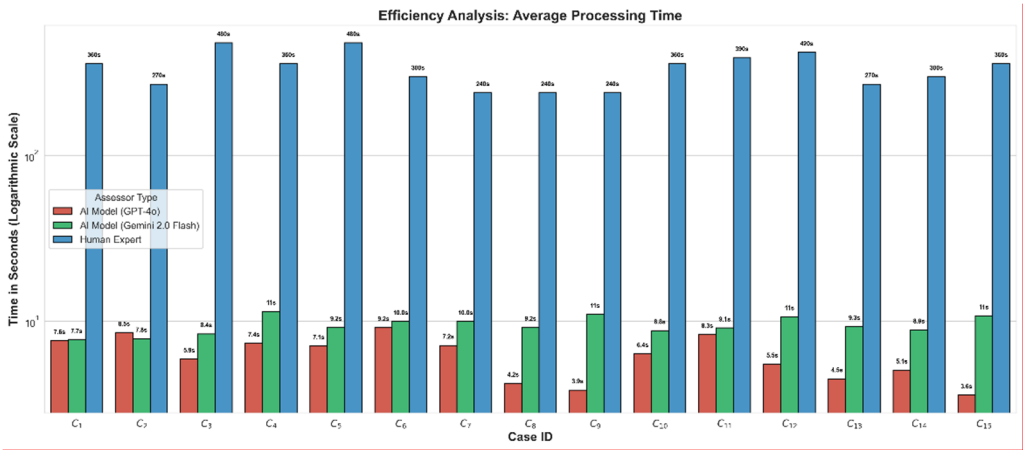
5.3.4 Operasyonel Verimlilik

Tek bir vakayı değerlendirmek için gereken işlem süresi insan uzmanları için yaklaşık 6 dakika (360 saniye) olarak ölçülmüştür. Buna karşın YZ ajanları, aynı görevi ortalama 4 saniyenin altında gerçekleştirmiştir (Bkz: Şekil 5.4). Bu durum, işlem hacminde yaklaşık 100 katlık bir artışa tekabül etmektedir. Özellikle, insan uzmanların değerlendirme öncesinde metrik şemasına zaten aşina oldukları unutulmamalıdır. Dolayısıyla raporlanan süre, öğrenme yükünden ziyade temel olarak analitik çabayı yansıtmaktadır. Elde edilen bu verimlilik artışı, ilk triyaj aşamasındaki insan darboğazını azaltarak, ölçeklenebilir ve gerçeğe yakın risk değerlendirmesini destekleyebilmektedir.

5.4 Niteliksel Fark Analizi ve Anlamsal Çıkarım Kapasitesi

Sayısal puanlamanın ötesinde YZ Ajanı, standart otomatik araçların genellikle gözden kaçırdığı bağlamsal nüansları yakalama yeteneğini

sergilemiştir. Tablo 5.4 ve Tablo 5.5'de ayrıntılı olarak açıklandığı üzere, Vaka C_1'de (Banka Hırsızlığı) YZ, "Ulusal Varlık Fonu" (Sovereign Funds) bağlamını tanıyarak risk seviyesini "Kritik" düzeye yükseltmiştir. Buna karşın, teknik açıklıkların önceliklendirilmesinde kullanılan standart çerçeveler, benzer mimari kusurları (tasarımsal zafiyetleri) genellikle "Orta" seviye olarak derecelendirmektedir. Bu gözlemler, RAG tarzı bağlam enjeksiyonunun, teknik zafiyet taraması ile iş odaklı risk yönetimi arasındaki boşluğu doldurmaya yardımcı olduğunu göstermektedir.



Şekil 5.6 Verimlilik Analizi: Ortalama İşlem Süresi (Logaritmik Ölçek)

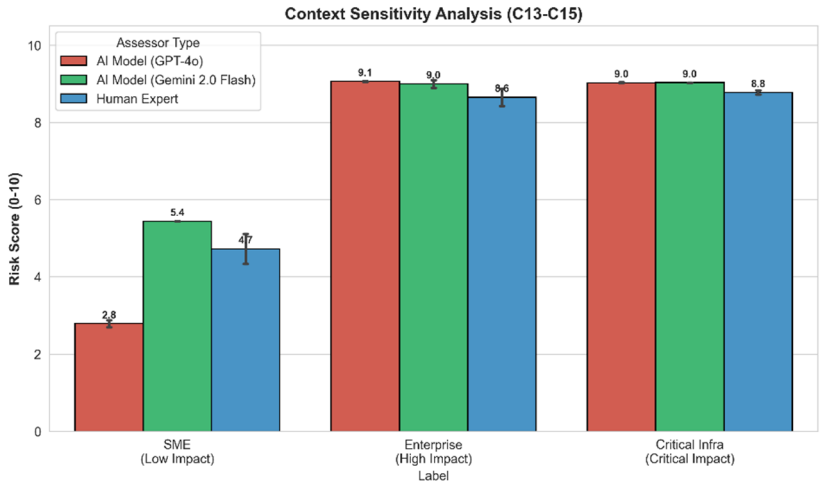
5.5 Kurumsal Bağlam Duyarlılığı ve Değişen Risk Algısı

Teknik şiddet ile iş etkisi arasındaki ayrımın yapılabilmesi modern risk değerlendirmesi için kritik bir gerekliliktir. LLM' in önerilen çerçeve içerisindeki bu kabiliyeti incelemek amacıyla; C_{13}, C_{14} ve C_{15} senaryoları dahil edilmiştir. Senaryolarda aynı teknik zafiyet (yama uygulanmamış RCE) kullanılmış, ancak farklı kurumsal bağlamlar enjekte edilmiştir.

- bir KOBİ tanıtım web sitesi,
- büyük bir işletmenin ödeme portalı ve
- kritik bir BT/OT ağgeçidi.

Şekil 5.5'te gösterildiği üzere çerçeve, risk skorlarını farlı varlık tanımlarına göre uyarlamıştır. KOBİ bağlamında (C_{13}) YZ ajanları; kritik verilerin ve yatay hareket yollarının bulunmamasını yansıtacak şekilde daha düşük skorlar üretmiştir (GPT-4o: 2,8; Gemini 2.0 Flash: 5,4). Aynı zafiyet bir kurumsal işletme (C_{14}) ya da bir kritik altyapı bağlamında (C_{15}) sunulduğunda, her iki model de yüksek varlık kritikliğine uygun olarak "Kritik" aralıkta (≈9,0) skorlar üretmiştir.

Genel olarak bu sonuçlar RAG tarzı bağlam enjeksiyonunun (YZ' nin sadece önceden öğretilen bilgilerle yetinilmeyip, her bir sorguda dışarıdan güvenilir kaynalar sunularak ilgii bilgilerin verilmesi), YZ ajanlarının kurumsal bağlamı puanlamaya dahil etmesine olanak tanıdığını göstermektedir. Bu sayede riskler, yalnızca statik teknik şiddet derecelendirmelerine dayanmak yerine iş etkisine göre de farklılaştırılabilmektedir.



Şekil 5.5 Bağlam Duyarlılığı Analizi: Üç farklı bağlamda (C13--C15)

5.6 Kontrol Önlemlerinin Etkisi ve Stres Testi Bulguları

Çerçevenin temel zafiyet puanlamasının ötesindeki mantıksal akıl yürütme yeteneğini değerlendirmek amacıyla; ShopFast Inc. Kurumsal ortamı (C_{14}) referans vaka olarak kullanılarak bir stres testi gerçekleştirilmiştir. Bu

kapsamda, teknik bulgunun sabit kaldığı ancak belirli telafi edici kontroller nedeniyle operasyonel bağlamın farklılaştığı, aynı yüksek şiddetli Log4Shell zafiyet senaryosunun üç farklı varyasyonu (C_{16} - C_{18}) sunulmuştur.

Tablo 5.7 Senaryolar ve Analiz Bulguları

Vaka No	Senaryo	Bulgular (Boşluk Analizi)
C_1	Banka Soygunu	YZ temsilcisi, 'Egemen Fonlar' bağlamını tanımlayarak risk seviyesini yükseltmiştir.
C_2	Savunma	YZ, mantık hatasını kurumun iç risk iştahıyla tutarlı bir şekilde modellemiştir.
C_3	Havacılık	YZ, endüstri standartlarıyla uyum sağlamış ve büyük ihlal olaylarını doğrulamıştır.
C_4	Sağlık	Gerçek dünyadaki aksamalar (NHS 111) ve düzenleyici para cezalarıyla doğrulanmıştır.
C_5	E-Ticaret	Yıkıcı doğasıyla tutarlı; kararlı küresel tehdit modellemesi sağlanmıştır.
C_6	SaaS	**Boşluk:** Statik puanların başarısız olduğu durumlarda yüksek ödül (bounty) miktarının etkiyi doğruladığı görülmüştür.
C_7	Siber Güvenlik	YZ, 'Platform Güveni' bağlamını tanımlayarak standart ölçütlerle uyum sağlamıştır.
C_8	Enerji / CNI	YZ, **"Faturalama Körlüğü"** paradoksunu doğru bir şekilde saptamıştır.
C_9	Tedarik Zinciri	İmzalı ikili dosyalardaki **"Güvenilir Hat"** (Trusted Pipeline) riskini tanımlamıştır.
C_{10}	Konaklama	Sesli kimlik avı (IdP baypas) üzerinden **"İnsan Odaklı Başarısızlığı"** tanımıştır.
C_{11}	Otomotiv	**"JIT Kırılganlığı"** modellenmiştir: 3. taraf ihlali ana hattı durdurmaktadır.
C_{12}	Telekom	**"Gölge API"** riskini tanımlamış ve "Test" etiketini geçersiz kılmıştır.

Tablo 5.8 YZ Puanlarının Gerçek Dünya Verileriyle Karşılaştırılması

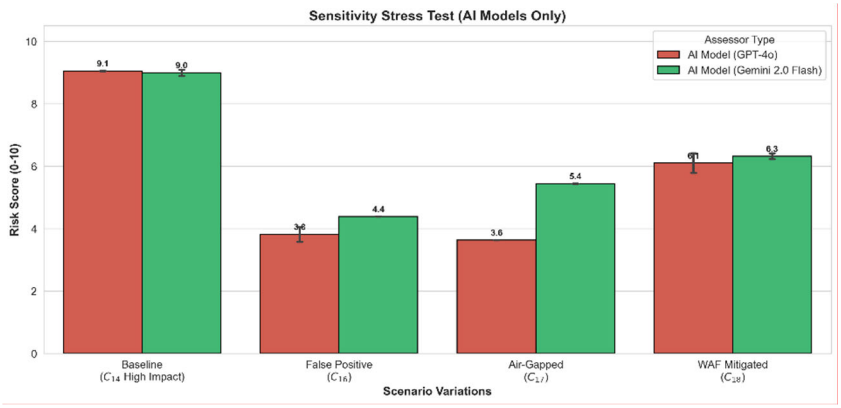
Vaka No	Teknik Şiddet	Etki Göstergesi	GPT-4o	Gemini
C_1	Uygulanamaz (Mimari Kusur)	81 Milyon $ Hırsızlık (NCSC, 2023)	8,93	9,16
C_2	Uygulanamaz (Mantık Hatası)	Pentest: Kritik (Vakalar Tablosu)	6,47	8,82
C_3	Kritik (9,4) (CVE-2023-4966)	İhlal: Boeing (CISA, 2023b)	8,5	9,04
C_4	Kritik (10,0) (CVE-2020-1472)	3,08 Milyon £ Ceza (ICO, 2021)	9,29	9,38
C_5	Kritik (9,8) (CVE-2024-34102)	Kitlesel Sömürü (4.000+ Mağaza) (Sansec, 2024)	8,65	9,19
C_6	Orta (4,8) (Satıcı Puanı)	3.500 $ Ödül (Yüksek Etki) (CrowdStrike, 2023)	7,48	8,61
C_7	Orta (5,3) (Satıcı Puanı)	2.500 $ Ödül (Orta Etki) (HackerOne, 2023)	4,49	5,27
C_8	Yüksek (7,0) (Bağımlılık eksik)	Ulusal Acil Durum (CISA, 2021)	9,26	9,26
C_9	Düşük / Güvenilir (Hatalı Negatif)	Küresel Casusluk (Sunburst) (CISA, 2020)	9,03	9,33
C_{10}	Uygulanamaz (Süreç Hatası)	Operasyonel Felç (SEC, 2023)	8,49	8,92
C_{11}	Düşük / Harici (3. taraf ihmali)	Üretim Durdurma (14 Fabrika) (Nikkei, 2022)	8,85	9,33
C_{12}	Orta (BOLA Atlanmış)	9,8 Milyon PII Sızıntısı (OAIC, 2022)	7,59	8,95

5.6.1 Telafi Edici Kontrollerin Risk Puanına Etkisi

Duyarlılık Stres Testi' nde (Bkz: Şekil 5.6) görüşlebileceği üzere YZ modelleri; yalnızca 10,0'lık CVSS temel puanına dayanmak yerine, telafi edici kontrollerin varlığına ve görünürdeki etkinliğine yanıt olarak risk puanlarını düşürmüştür.

- **Kurumsal Temel Çizgi (C_{14}):** Herhangi bir iyileştirme yapılmadığı durumda YZ ajanları; korumasız bir Ödeme Geçidinin "Marka Güveni" ve "PCI DSS" üzerindeki etkilerine atıfta bulunarak riski "Kritik" aralıkta (GPT-4o: 9,1; Gemini: 9,0) değerlendirmiştir.
- **Yanlış Pozitif (C_{16}) (False-Positive):** Bağlamda, zafiyet tarayıcının "yürütülebilir olmayan bir dosyayı" işaret ettiğini ve zafiyet barındıran kütüphanenin belleğe yüklenmediğini belirtilmiştir. Bu durumda YZ modelleri puanı "Düşük" aralığa (GPT-4o: 3,8 ; Gemini: 4,4) indirmiştir. Özellikle puanın 0,00'a düşmemesi; triyajın operasyonel maliyetini ve ikili statik tarayıcılar tarafından genellikle göz ardı edilen bir nüans olan hatalı teşhis olasılığından kaynaklanan artık riski yansıtmaktadır.
- **İnternet Bağlantısı Olmayan / Çevrimdışı (C_{17}):** Güvenli bir ortamda fiziksel olarak bağlantısı kesilmiş yedekleme sunucusu için puan 3,6--5,4 aralığına gerilemiştir. YZ çıktıları; yazılım kusurunun varlığını sürdürmesine rağmen, ağ bağlantısının bulunmamasının uygulanabilir saldırı vektörlerini kısıtladığını, C_{14}' e göre belirgin bir skor düşüşü ile göstermektedir. Fakat fiziksel erişimin gerçekleşmesi halinde hala daha istismar edilebilir bir zafiyetin bulunmasını da artık risk değeri ile göstermektedir.
- **WAF İyileştirmesi (C_{18}):** Bir Bulut WAF'ın "ENGELLEME" (BLOCK) modunda aktif olduğu senaryoda risk, "Orta/Yüksek" aralıkta (GPT-4o: 6,1; Gemini: 6,3) çıkmıştır. Çıktılar, bu durumu tamamen "Düzeltilmiş" olarak ele almak yerine WAF'ların atlatılabileceği (bypass) gerçeğini yansıtmış; durumu "Giderilmiş" (Remediated) yerine "İyileştirilmiş" (Mitigated) olarak sınıflandırarak gerçekçi bir yaklaşım sergilemiştir.

Genel olarak bu farklılaştırılmış puanlama; çerçevenin teorik zafiyet şiddeti (CVSS) ile bağlama dayalı iş riski arasında ayrım yaparak "Artık Riski" (Residual Risk) operasyonelleştirdiğine dair kanıt sunmaktadır.



Şekil 5.9 Duyarlılık Stres Testi (C14-C16-C17-C18 kıyaslaması)

BÖLÜM 6

6. TARTIŞMA VE DEĞERLENDİRME

Bu çalışma sonucunda elde edilen veriler otomatik risk değerlendirmesinin karmaşık ve "kara kutu" (black-box) derin öğrenme modelleri gerektirdiği yönündeki yaygın varsayımın aksini savunmaktadır. Elde edilen bulgular, uzmanlarca doğrulanmış metriklerin ve kamuya açık LLM'lerin anlamsal akıl yürütme (*semantic reasoning*) kabiliyetlerinin etkin kullanılabileceğini göstermektedir. Bu yaklaşım sayesinde, zafiyetlerin bağlamsallaştırılması gibi üst düzey bilişsel görevlerin, hız ve hassasiyet açısından önemli kazanımlarla otomatikleştirilebileceği ortaya konmaktadır. Bununla birlikte, geliştirme sürecinde LLM'lerde tekrarlayan çeşitli davranışsal yapaylıklar (behavioral artifacts) gözlemlenmiş; ancak bunlar hedefe yönelik istem mühendisliği (prompt engineering) stratejileri aracılığıyla giderilmiştir.

6.1 Algoritmik Sapmalar ve İyileştirme Stratejileri

Çerçevenin yinelemeli geliştirme sürecinde, tekrarlanan iki davranışsal örüntü tespit edilmiş ve ele alınmıştır. Söz konusu gözlemler, risk değerlendirme süreçlerinde LLM'lerin operasyonel özelliklerine ilişkin pratik içgörüler sunmaktadır.

6.1.1 Katmanlı Ayrıştırma ile Sorumlulukların Ayrılması

Erken prototip aşamasında, yönetişim kuralları (Örn: "Eğer bağlam eksikse, Yüksek seçeneğini belirle") doğrudan JSON Çıktı Şeması tanımlarına dahil edilmiştir. Bu durumun, modellerin şemaya gömülü mantıksal kısıtlamaları karşılamak adına asılsız tehdit çıkarımları üretmesine yol açan ve "İstem Sızıntısı" (Instruction Leakage) olarak adlandırılan bir fenomene katkıda bulunduğu gözlemlenmiştir.

Söz konusu sorunu gidermek amacıyla, nihai mimaride "Sorumlulukların Ayrılması" (Separation of Concerns) tasarımı benimsenmiştir (Bkz: Ek A). JSON Şeması yalnızca veri yapılandırma tanımlarıyla sınırlandırılmış (Örn: *"insufficient_information": false* , Bkz: Şekil A.1). Akıl yürütme mantığı ise Sistem İstemi' ne (System Prompt) taşınarak dinamik mantık enjeksiyonu yoluyla uygulanmıştır (Bkz: Şekil A.1). Bu ayrıştırma, şema kaynaklı halüsinasyonları azaltmış ve LLM tabanlı risk motorlarının kararlılığını artırmak için pratik bir tasarım yaklaşımı sağlamıştır.

6.1.2 "Önce Güvenlik" Odaklı Sistematik Sapmanın Analizi

Optimize edilmiş isteme rağmen yapılan analizler; insan medyanına kıyasla sistematik bir pozitif sapma olarak tanımlanan, kasıtlı bir mimari "Önce Güvenlik Yanlılığına" (Safety-First Bias) işaret etmektedir. Şekil 5.1'de gösterildiği üzere YZ ajanları, 15 senaryonun ortalama 11,5'inde (Gemini 2.0 Flash için 14'e kadar) daha yüksek risk skorları üretmiştir; bu durum insan temel çizgisinin ortalama %2,059 (0,2059 puan) üzerinde bir kaymaya tekabül etmektedir. Bu davranış, teknik bulgunun mevcut olduğu ancak etkili bir şekilde nötralize edildiği (yürütülebilir olmayan dosya) Yanlış Pozitif senaryosunda (C_{16}) belirgin şekilde görülmektedir.

Her iki LLM' yanıtında da risk skoru 0,0'a indirmek yerine, "Orta" aralıkta bir "Artık Risk" skorunu (GPT-4o: 3,86 ; Gemini 2.0: 4,39) korumuştur. Bu durum teorik bir sıfır değerine göre aşırı bir tahmin gibi görünse de siber güvenlik riskindeki maliyet asimetrisi ile tutarlıdır. Tespit edilemeyen bir ihlal nedeniyle hatalı negatifin (false-negative) etkisi ağır olabilirken; hatalı pozitif (false-positive) durumunda temel olarak manuel triyaj için operasyonel ek yük oluşmaktadır. Bu ihtiyatlı tutum operasyonel maliyeti, gürültü azaltımının önüne koyarak güvenlik payını (safety margin) artırmaktadır.

6.2 Bağlamsal Duyarlılık ve Matematiksel Risk Alt Sınırı

Elde edilen sonuçlar "Sıfır Risk" skorunun (0,00), metrik şeması tarafından tanımlanan operasyonel sınırlar dahilinde ulaşılamaz olduğunu göstermektedir. C_{14} (Temel Çizgi) ve C_{16} (Hatalı Pozitif) arasındaki karşılaştırmada gözlemlendiği üzere; puan 9,06'dan 3,86'ya düşmüş ancak sıfıra ulaşmamıştır. Bu artık skor; "Sektörel Duyarlılık" ve "Yama Durumu" gibi, sıfır olmayan operasyonel alt sınırlara sahip parametrelerin katkısını yansıtmaktadır.

Benzer şekilde, C_{14} senaryosundan C_{17} (İzole/Çevrimdışı) vakasına geçiş, modelin tanım tabanlı mantık kullanımını örneklendirmektedir. Ajan, yalnızca zafiyet raporuna dayanmak yerine, izolasyon bağlamıyla tutarlı olacak şekilde Saldırı Vektörünü "Ağ" üzerinden "Fiziksel" olarak yeniden sınıflandırmıştır. C_{14} (9,06) ile C_{13} (KOBİ bağlamı, 2,73) arasındaki zıtlık; modelin "Marka Değeri" ve "İş Etkisi" gibi kurumsal faktörleri puanlamaya dahil ettiğini göstermektedir. Böylece yalnızca bir hesaplayıcı olarak değil, aynı zamanda bağlama duyarlı bir değerlendirici olarak hareket ettiğine işaret etmektedir.

6.3 Yüksek Hassasiyetli ve Tekrarlanabilir Bulgular

Bu çalışmanın temel bulgularından biri, geliştirilen çerçevenin kontrollü koşullar altında olasılıksal modellerden oldukça kararlı çıktılar elde edebilmesidir. Çerçeve mimarisi (Bkz: Şekil 4.1) çok kaynaklı veri entegrasyonunu destekleyecek şekilde tasarlanmış olsa da, bu çalışmadaki deneysel odak kasten zafiyet ve kurumsal girdiler ile sınırlandırılmıştır. Bilimsel titizliği odağa alan ve dış değişkenleri minimize eden bu stratejik kapsam daraltması, sonuçların yüksek sadakatle tekrarlanabilir olmasını doğrudan mümkün kılmıştır. LLM eleştirmenlerince; denetim uyumlu ortamlarda deterministik olmama durumu, genellikle bu modellerin benimsenmesinin önünde bir engel olarak nitelendirilmektedir.

Buna karşın ampirik sonuçlar (Bkz: Tablo 5.3), bu kısıtlamanın önerilen mimari dahilinde önemli ölçüde hafifletildiğine işaret etmektedir. Elde

edilen yüksek seçim uyumu (High Selection Agreement) (>%97), modellerin operasyonel kararlılığını kanıtlamaktadır. Söz konusu istatistiksel tutarlılık, yapılandırılmamış girdilerin uzmanlarca doğrulanmış katı bir şema (Metrik Kutusu) ile eşleştirilmesi yoluyla sağlanmıştır. Bu bulgu; kamuya açık LLM'lerin anlamsal yorumlama yeteneklerini muhafaza ederken, statik algoritmalarla karşılaştırılabilir bir disiplinle davranabileceğini göstermektedir. Bu durum ise istikrarsızlığın, modellerin doğasından ziyade genellikle kısıtlanmamış istemlerle (unconstrained prompting) ilişkili olduğunu göstermektedir.

6.4 Bulguların Yorunlanması: Hız ve Doğruluk

Değerlendirme döngü süresinde gözlemlenen yaklaşık 100 katlık (~100X) iyileşme, risk yönetimindeki temel darboğazın veri mevcudiyeti değil, verilerin sentezlenme süreci olduğunu göstermektedir. İnsan analistlerin teknik zafiyetler, varlık profilleri ve kurumsal bağlam arasındaki etkileşimi değerlendirmek için ortalama altı dakika (360 saniye) harcaması, buna karşın LLM ajanlarının sentez sürecini ortalama dört saniyenin altında tamamlaması ciddi ölçüde hızlanma olduğunu göstermektedir.

Önemli bir nokta da elde edilen bu verimlilik artışı insan temel çizgisiyle olan uyumda bir kayba yol açmamasıdır. İnsan medyanına kıyasla elde edilen düşük değişkenlik (Örn: $\sigma < 0,1$) ve güçlü korelasyon (Örn: $r \approx 0,97$), "Metrik Kutusu" yapısının etkili bir koruma sınırı işlevi gördüğünü kanıtlamaktadır. Toplu olarak değerlendirildiğinde bu sonuçlar, ajanın genel insan puanlama eğilimlerini takip ettiğini ortaya koymaktadır. Bununla birlikte sistem, risk asimetrisini dengelemek adına daha ihtiyatlı bir eşikte çalışarak, daha önce tartışılan güvenlik öncelikli mimari yanlılığını da korumaktadır.

6.5 Sistem Şeffaflığında Uzman Doğrulamasının Payı

Eleştirmenler, ticari LLM'lere olan bağımlılığın şeffaf olmayan (opak) bir yapı oluşturduğunu savunabilmektedirler. Bu endişe; akıl yürütme (reasoning) aşamasının, puanlama (scoring) aşamasından ayrıştırılmasıyla giderilmiştir. Uçtan uca derin öğrenme yaklaşımlarında kararın gerekçesi genellikle örtük (latent) kalmaktadır. Geliştirilen çerçeve ise YZ ajanının tüm seçimlerini gerekçelendirmesini zorunlu kılarak bu kapalılığı ortadan kaldırmaktadır. Bu süreçte ajan, yapılandırılmamış bağlam verilerini insan tarafından okunabilen belirli alt parametrelerle eşleştirerek kararlarını şeffaf hale getirmektedir. LLM'ler kapalı kaynaklı bir yapıya sahip olsa da, üretilen risk skorları tamamen izlenebilir bir matematiksel temele dayanmaktadır. Nihai puanlar, doğrudan uzman grubundan elde edilen ROC ağırlıkları üzerinden hesaplandığı için her adımın doğrulanması mümkündür. Bu şeffaf mimari, siber güvenlik denetimleri açısından kritik bir gereklilik olan süreç şeffaflığını doğrudan desteklemektedir.

6.6 Çalışmanın Kısıtlılıkları

Başarılı sonuçlara rağmen, çalışmanın belirli kısıtlılıkları mevcuttur. Birincisi; çerçeve kısmen LLM'nin temel bilgi birikimine dayanmaktadır. Bağlam enjeksiyonu güncel veriler sağlasa da, modelin eğitim veri kesim tarihi (training cut-off), yeni ortaya çıkan (örneğin sıfırıncı gün/zero-day) tekniklerin işlenmesini sınırlayabilmektedir. İkincisi; İhtiyatlı Tahmin Protokolü, belirsiz vakalarda hafifçe yükseltilmiş skorlar üretebilmekte ve bu durum Risk Sorumlusu tarafından manuel düzeltme gerektirebilmektedir.

Üçüncüsü; Tablo 3.1'de raporlanan ROC kaynaklı ağırlıklar, evrensel bir sabit yerine uzman odaklı bir başlangıç (temel çizgi) olarak yorumlanmalıdır. Söz konusu ağırlıklar 101 profesyonelin görüşlerine dayansa da bölgesel veya sektörel bakış açılarını yansıtabilmektedir. Bu nedenle çerçeve, periyodik yeniden kalibrasyonu desteklemekte ve Bölüm 3.5.2'de detaylandırılan Eşitlik 3.2 ve Eşitlik 3.3 ile açıklandığı üzere, kurumların ağırlıkları kendi risk

iştahlarına göre özelleştirmelerine olanak tanımaktadır. Son olarak, sistemin uzun vadeli kararlılığını değerlendirmek için canlı bir SOC (Güvenlik Operasyon Merkezi) ortamında boylamsal testlere ihtiyaç duyulmaktadır.

SONUÇ VE ÖNERİLER

Statik zafiyet verilerini kurumsal bağlamla gerçeğe yakın zamanlı olarak birleştiren bu çalışma; uzman bilgisini operasyonelleştiren, siber risk yönetiminde statik yaklasımlardan bağlama duyarlı akıl yürütmeye geçişi destekleyen bir çerçeve önermektedir.

Bu çalışmanın temel katkıları ve bulguları aşağıda özetlenmiştir:

- **Uzman Onaylı Metodoloji:** 101 siber güvenlik profesyonelinden alınan sıra tabanlı derecelendirmeleri, ROC yöntemini kullanarak nesnel parametre ağırlıklarına dönüştüren bir ağırlık belirleme metodolojisi sunulmuştur.
- **Mimari İnovasyon (Sorumlulukların Ayrılması):** Mantıksal kuralların doğrudan çıktı şemasına gömülmesinin model halüsinasyonlarına katkıda bulunduğu ve "İstem Sızıntısı" (Instruction Leakage) olarak adlandırılan fenomen tespit edilmiştir. Mantık Katmanı (Sistem İstemi) ile Tanım Katmanı'nı (JSON Şeması) birbirinden ayıran "sorumlulukların ayrılması" tasarımı benimsenerek bu sorun giderilmiştir.
- **Davranışsal Analiz ("Önce Güvenlik"):** Yapılan analizler, insan medyanına kıyasla sistematik bir pozitif sapma olarak tanımlanan mimari bir "Önce Güvenlik Yanlılığına" işaret etmiştir. Nicel sonuçlar, bu örüntünün 15 senaryonun ortalama 11,5'inde görüldüğünü ve insan temel çizgisinin ortalama %2,059 (0,2059 puan) üzerinde bir kayma yarattığını göstermiştir. Bu kayma, hatalı negatifleri (false negative) azaltmayı, hatalı pozitiflerin (false positive) operasyonel yükünü minimize etmeye tercih eden bir risk yönetimi duruşuyla tutarlıdır. Çerçeve, stres testlerinde (C_{16}) tanım tabanlı mantığı kullanarak artık riski yakalamış; böylece sayısal puanlama ile operasyonel güvenlik mülahazaları arasında bir denge kurulmasını desteklemiştir.

- **Performans ve Doğrulama:** Çerçeve, değerlendirme döngü süresini 100 kattan fazla azaltırken, insan medyanıyla güçlü bir uyum (Pearson *r* değeri 0,9390 ile 0,9717 aralığında) sergilemiştir. Gerçek dünya vakaları (Örn: Colonial Pipeline, Optus) üzerindeki karşılaştırmalı değerlendirmeler, sistemin teknik şiddet ile iş etkisi arasındaki ayrımı yapabildiğini göstermiştir.

Bu araştırmanın gelecekteki genişletmeleri iki yöne odaklanacaktır:

1. **Gizliliği Koruyan Yerel Modeller:** Regülasyonlara tabii sektörlerde veri güvenliği gereksinimlerini karşılamak amacıyla, genel amaçlı kamuya açık LLM'lerden yerel olarak konuşlandırılan ince ayarlı Küçük Dil Modellerine (SLM) (Örn: Llama-3 8B) geçiş yapılması.
2. **Otomatik Risk İyileştirme (SOAR Entegrasyonu**): Çerçevenin, belirlenen risk profiline göre makine tarafından okunabilir iyileştirme yönergeleri (Örn: Ansible/Terraform betikleri) üretecek şekilde genişletilmesi ve böylece siber savunma iş akışlarının otomasyonunun artırılması.

Bu genişletmelerin hayata geçirilmesiyle çerçevenin gerçek zamanlı, YZ destekli, denetlenebilir ve şeffaf bir risk yönetişimi için kurumsal kullanıma hazır bir çözüme dönüşmesi hedeflenmektedir.

VERİ PAYLAŞIM BEYANI

Bu çalışmanın bulgularını destekleyen veriler (doğrulama senaryoları, ham puanlama günlükleri ve tam JSON şeması dahil), makul bir talep üzerine ilgili yazardan temin edilebilir.

KAYNAKLAR

Abbas, G., Ali, M., Ahmad, M. ve Khan, A. (2025). CIRA-Cyber intelligent risk assessment methodology for industrial internet of things based on machine learning. IEEE Access, 13, 77001–77016. https://doi.org/10.1109/ACCESS.2025.3559617

Ahmed, M. G., Panda, S., Xenakis, C. ve Panaousis, E. (2022). MITRE ATT&CK-driven cyber risk assessment. Proc. 17th International Conference on Availability, Reliability and Security (ARES 2022) içinde (ss. 1–10). New York, NY, USA. https://doi.org/10.1145/3538969.3544420

Aksu, M. U., Dilek, M. H., Tatlı, E. İ., Bicakci, K., Dirik, H. İ., Demirezen, M. U. ve Aykır, T. (2017). A quantitative CVSS-based cyber security risk assessment methodology for IT systems. Proc. 2017 IEEE International Carnahan Conference on Security Technology (ICCST) içinde (ss. 1–8). https://doi.org/10.1109/CCST.2017.8167819

Alberts, C. ve Dorofee, A. (2002). Managing information security risks: The OCTAVE approach. Addison-Wesley Professional.

Al-Sada, B., Sadighian, A. ve Oligeri, G. (2024). Analysis and characterization of cyber threats leveraging the MITRE ATT&CK database. IEEE Access, 12. https://doi.org/10.1109/ACCESS.2023.3344680

Bansal, U., Sikka, G., Awasthi, L. K. ve Bhargava, B. (2024). Quantitative evaluation of extensive vulnerability set using cost benefit analysis. IEEE Transactions on Dependable and Secure Computing, 21(1), 298–308. https://doi.org/10.1109/TDSC.2023.3253121

Blount, J. (2021). Cyber threats in the pipeline: Lessons from the federal response to the colonial pipeline ransomware attack. U.S. House Committee on Homeland Security (117. Kongre) Önünde İfade.

Camacho, J. M., Couce-Vieira, A., Arroyo, D. ve Insua, D. R. (2025). A cybersecurity risk analysis framework for systems with artificial intelligence components. International Transactions in Operational Research. https://doi.org/10.1111/itor.70049

Cheimonidis, P. ve Rantos, K. (2025). A proactive and time-sensitive cyber risk assessment model integrating markov chains and bayesian networks. IEEE Access, 13. https://doi.org/10.1109/ACCESS.2025.3575070

Cimpanu, C. (2022). Toyota halts production after suspected cyberattack at supplier. The Record by Recorded Future. https://therecord.media/toyota-halts-production-after-suspected-cyberattack-at-supplier adresinden 29 Kasım 2025 tarihinde erişildi.

Citrix Systems Inc. (2023). Citrix bleed vulnerability and advanced data theft: CVE-2023-4966. https://nvd.nist.gov/vuln/detail/cve-2023-4966 adresinden 29 Kasım 2025 tarihinde erişildi.

CyberArk. (2023). The MGM resorts attack: Analysis of the identity-based compromise. CyberArk Threat Research. https://www.cyberark.com/resources/blog/the-mgm-resorts-attack-initial-analysis adresinden 29 Kasım 2025 tarihinde erişildi.

Cybersecurity and Infrastructure Security Agency (CISA). (2020). Emergency directive 21-01: Mitigate SolarWinds Orion code compromise. U.S. Department of Homeland Security. https://www.cisa.gov/news-events/directives/ed-21-01-mitigate-solarwinds-orion-code-compromise adresinden 29 Kasım 2025 tarihinde erişildi.

Cybersecurity and Infrastructure Security Agency (CISA). (2023). #StopRansomware: LockBit 3.0 ransomware affiliates exploit CVE 2023-4966 Citrix bleed vulnerability. Alert AA23-325A. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-325a adresinden 29 Kasım 2025 tarihinde erişildi.

CISA ve FBI. (2021). DarkSide ransomware: Best practices for preventing business disruption from ransomware attacks. Cybersecurity and Infrastructure Security Agency, Alert AA21-131A. https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-131a adresinden 29 Kasım 2025 tarihinde erişildi.

Drissi, S., Chergui, M. ve Khatar, Z. (2025). A systematic literature review on risk assessment in cloud computing: Recent research advancements. IEEE Access, 13. https://doi.org/10.1109/ACCESS.2025.3561123

Femi, A. G. ve Madu, M. (2025). Enhancing adaptive cybersecurity risk management through AI-driven threat detection. International Journal of Trendy Research in Engineering and Technology, 9(2). https://doi.org/10.54473/IJTRET.2025.9210

Forum of Incident Response and Security Teams (FIRST). (2023). Common vulnerability scoring system (CVSS) v4.0: Specification document.

https://www.first.org/cvss/v4.0/specification-document adresinden erişildi.

Freund, J. A. ve Jones, J. (2014). Measuring and managing information risk: A FAIR approach. Butterworth-Heinemann.

HackerOne. (2025a). Unauthorized partner access via invitation process - Report 2885269. HackerOne Security Report. https://hackerone.com/reports/2885269 adresinden 29 Kasım 2025 tarihinde erişildi.

HackerOne. (2025b). Business logic flaw allowing self-creation of testimonials for reputation manipulation - Report 2490953. HackerOne Security Report. https://hackerone.com/reports/2490953 adresinden 29 Kasım 2025 tarihinde erişildi.

He, W., Li, H. ve Li, J. (2019). Unknown vulnerability risk assessment based on directed graph models: A survey. IEEE Access, 7. https://doi.org/10.1109/ACCESS.2019.2954092

Hmimou, Y., Tabaâ, M., Khiat, A. ve Hidila, Z. (2025). A multi-agent system for cybersecurity threat detection and correlation using large language models. IEEE Access, 13. https://doi.org/10.1109/ACCESS.2025.3602681

Information Commissioner's Office (ICO). (2025). Advanced penalty notice - Advanced Computer Software Group, 26 March 2025. UK Government Publication. https://ico.org.uk/media2/gdlfddgc/advanced-penalty-notice-20250327.pdf adresinden 29 Kasım 2025 tarihinde erişildi.

International Organization for Standardization. (2018). ISO 31000:2018 - Risk management — Guidelines.

International Organization for Standardization. (2022). ISO/IEC 27005:2022 - Information security, cybersecurity and privacy protection — Guidance on managing information security risks.

Islam, S., Basheer, N., Papastergiou, S., Ciampi, M. ve Silvestri, S. (2025). Intelligent dynamic cybersecurity risk management framework with explainability and interpretability of AI models for enhancing security and resilience of digital infrastructure. Journal of Reliable Intelligent Environments. https://doi.org/10.1007/s40860-025-00253-3

Jamili, L. K., Bhanuvardhan, B., Rawat, H., Semlani, D., Garg, V. ve Goel, O. (2025). Artificial intelligence for adaptive risk assessment in cloud-based security frameworks. Proc. 2025 International Conference on Networks

and Cryptology (NETCRYPT) içinde. https://doi.org/10.1109/NETCRYPT65877.2025.11102751

Kawanishi, Y., Nishihara, H., Yoshida, H., Yamamoto, H. ve Inoue, H. (2023). A study on threat analysis and risk assessment based on the 'Asset Container' method and CWSS. IEEE Access, 11, 18148–18156. https://doi.org/10.1109/ACCESS.2023.3246497

Malik, A., Arshid, K., Noonari, N. ve Munir, R. (2025). Artificial intelligence-driven cybersecurity framework using machine learning for advanced threat detection and prevention. Scholars Journal of Engineering and Technology, 13(6), 401–423. https://doi.org/10.36347/sjet.2025.v13i06.005

Marinho, R. ve Holanda, R. (2023). Automated emerging cyber threat identification and profiling based on natural language processing. IEEE Access, 11. https://doi.org/10.1109/ACCESS.2023.3260020

MGM Resorts International. (2023). Form 8-K: Report of unscheduled material events or corporate event. U.S. Securities and Exchange Commission (SEC). https://www.sec.gov/ix?doc=/Archives/edgar/data/789570/000119312523251667/d461062d8k.htm adresinden 29 Kasım 2025 tarihinde erişildi.

Microsoft Security Response Center. (2020). Deep dive into the solorigate 2nd stage activation: From SUNBURST to TEARDROP. Microsoft Security Blog. https://www.microsoft.com/en-us/security/blog/2020/12/18/analyzing-solorigate-the-compromised-dll-file-that-started-a-sophisticated-cyberattack-and-how-microsoft-defender-helps-protect/ adresinden 29 Kasım 2025 tarihinde erişildi.

MITRE Corporation. (2011). Common Weakness Scoring System (CWSS). https://cwe.mitre.org/cwss/ adresinden erişildi.

National Institute of Standards and Technology (NIST). (2012). Guide for conducting risk assessments. SP 800-30 Rev. 1.

Nikkei Asia. (2022). Toyota to suspend all Japan factory operations March 1 after cyberattack. Nikkei Asia Review. https://asia.nikkei.com/Business/Automobiles/Toyota-to-suspend-all-Japan-factory-operations-March-1-after-cyberattack adresinden 29 Kasım 2025 tarihinde erişildi.

Office of the Australian Information Commissioner (OAIC). (2022). OAIC opens investigation into Optus over data breach. Australian Government. https://www.oaic.gov.au/news/media-centre/oaic-opens-investigation-into-optus-over-data-breach adresinden 29 Kasım 2025 tarihinde erişildi.

Optus. (2022). Optus notifies customers of cyberattack compromising customer information. Optus Media Centre. https://www.optus.com.au/about/media-centre/media-releases/2022/09/optus-notifies-customers-of-cyberattack adresinden 29 Kasım 2025 tarihinde erişildi.

Sansec Research Team. (2025). CosmicSting attack & defense overview. Sansec. https://sansec.io/research/cosmicsting adresinden 29 Kasım 2025 tarihinde erişildi.

UK National Cyber Security Centre. (2025). Nation-state hackers case study: Bangladesh Bank heist. https://cyber.uk/areas-of-cyber-security/cyber-security-threat-groups-2/nation-state-hackers-case-study-bangladesh-bank-heist/ adresinden 29 Kasım 2025 tarihinde erişildi.

Unal, N. M. ve Celiktas, B. (2025). A metric-driven IT risk scoring framework: Incorporating contextual and organizational factors. Proc. 2025 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA) içinde. Antalya, Türkiye. https://doi.org/10.1109/ACDSA65407.2025.11166074

Wang, L., Ali, Y., Nazir, S. ve Niazi, M. (2020). ISA evaluation framework for security of internet of health things system using AHP–TOPSIS methods. IEEE Access, 8. https://doi.org/10.1109/ACCESS.2020.3017221

Wang, T., Lv, Q., Hu, B. ve Sun, D. (2020). CVSS-based multi-factor dynamic risk assessment model for network system. Proc. 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC) içinde (ss. 289–294). https://doi.org/10.1109/ICEIEC49280.2020.9152340

Wang, W., Shi, F., Zhang, M., Xu, C. ve Zheng, J. (2020). A vulnerability risk assessment method based on heterogeneous information network. IEEE Access, 8, 148315–148330. https://doi.org/10.1109/ACCESS.2020.3015551

Xie, W., Yu, X., Zhang, Y. ve Wang, H. (2025). An improved shapley value benefit distribution mechanism in cooperative game of cyber threat intelligence sharing. University of Chinese Academy of Sciences, Beijing, China.

Yang, H., Yuan, H. ve Zhang, L. (2023). Risk assessment method of IoT host based on attack graph. Mobile Netw. Appl. https://doi.org/10.1007/s11036-023-02198-4

Younang, V. C. W. ve Sen, A. (2025). Security risk assessment using bayesian attack graphs and complex probabilities for large scale IoT applications.

IEEE Transactions on Dependable and Secure Computing, 22(6), 7360–7371. https://doi.org/10.1109/TDSC.2025.3597186

Yu, J., Shvetsov, A. V. ve Alsamhi, S. H. (2024). Leveraging machine learning for cybersecurity resilience in industry 4.0: Challenges and future directions. IEEE Access, 12. https://doi.org/10.1109/ACCESS.2024.3482987

Zhang, Y., Wang, Z., Wang, Y., Lin, K., Li, T., Liu, H., Li, C. ve Wang, B. (2023). A risk assessment model for similar attack scenarios in industrial control systems. The Journal of Supercomputing, 79, 15955–15979. https://doi.org/10.1007/s11227-023-05269-1

EKLER

EK A: TEKRAR ÜRETİLEBİLİRLİK VE SİSTEM BİLEŞENLERİ

Önerilen deterministik risk puanlama metodolojisinin şeffaflığını ve tekrar üretilebilirliğini desteklemek amacıyla; deneylerde kullanılan temel sistem talimatları, çıktı şeması ve bir gerçek dünya vaka çalışması bu bölümde sunulmaktadır.

A.1 Temsilci Talimat Seti (Sistem İstemi)

Aşağıdaki talimat seti, LLM için temel direktifi sağlamakta ve Şekil Ek A.1'de tanımlanan JSON çıktı formatını zorunlu kılmaktadır.

- **Sistem Kimliği ve Amacı:** Temsilci, "Kıdemli Siber Risk Analisti" rolünde tanımlanmıştır.
- **Hedef:** Yalnızca sağlanan girdi verilerine dayanarak bilimsel olarak titiz bir risk değerlendirmesi yapmaktır.
- **Çıktı Formatı:** Temsilcinin kesinlikle geçerli bir JSON nesnesi çıktısı üretmesi zorunludur.

A.2 Mantık Kuralları Enjeksiyonu

"Güvenlik Öncelikli" (Safety-First) tasarımı desteklemek ve belirsiz bağlamlarda halüsinasyon riskini azaltmak amacıyla, Şekil A.2' deki kural seti çalışma zamanında kullanıcı istemine dinamik olarak enjekte edilmektedir. Bu geçersiz kılmalar (overrides), belirli uç durumlar (örneğin, hava boşluklu varlıklar) için deterministik bir koruma işlevi görerek LLM'in olasılıksal çıktılarını sınırlandırmaktadır.

A.3 JSON Çıktı Şeması (Metrik Kutusu)

Otomatik ayrıştırma ve istatistiksel analizi kolaylaştırmak amacıyla temsilci, Tablo 3'te tanımlanan "Metrik Kutusu"na karşılık gelen sabit bir JSON yapısıyla sınırlandırılmıştır. Şekil Ek A.3, ağırlıkların (WF) ve puanlama seçeneklerinin hiyerarşisini gösteren kısaltılmış bir şemayı sunmaktadır9.

SYSTEM IDENTITY & PURPOSE
ROLE: You are a Senior Cyber Risk Analyst.
GOAL: Perform a scientifically rigorous risk assessment based ONLY on the provided input data.

OUTPUT FORMAT
You must output a strictly valid JSON object. Do not include markdown fences (```json) or introductory text.

JSON STRUCTURE:

```
{
  "asset_info": {
    "name": "Extract from context",
    "id": "Extract from context"
  },
  "selected_metrics": {
    // Map EVERY parameter from the input Schema to one of its exact Options.
    // Example: "Attack Vector": "Network"
  },
  "parameter_explanations": {
    "Attack Vector": {
      "reason": "Brief justification based on context.",
      "insufficient_information": false
    }
  }
}
```

Şekil A.1 JSON yapısını zorunlu kılan sistem istemi

CRITICAL LOGIC OVERRIDES

You must analyze the <organization_context> and apply the following overrides STRICTLY.
These rules serve to map compensating controls to their functional risk-equivalent states within the framework, superseding technical severity values to reflect baseline residual risk.

--- RULE 1: OFFLINE / AIR-GAPPED ASSETS ---
IF the context mentions "Air-Gapped", "Offline", "Disconnected", "No Internet", or "Cold Storage":
1. SELECT "Physical" for "Attack Vector". (Reason: Remote attack is impossible)
2. SELECT "Fully Patched" for "Patch Availability". (Reason: Physical isolation functionally neutralizes the remote vulnerability, reaching the framework's baseline residual risk score of 0.25)
3. SELECT "None" for "Confidentiality Impact".
4. SELECT "None" for "Integrity Impact".
5. SELECT "None" for "Availability Impact".
6. SELECT "No impact" for "Technical Impact".

--- RULE 2: FALSE POSITIVE ---
IF the context mentions "False Positive", "Invalid Alert", "Investigation Closed", or "No Risk":
1. SELECT "None" or "No impact" for ALL impact metrics.
2. SELECT "Fully Patched" for "Patch Availability". (Reason: Non-existent or non-exploitable vulnerabilities are mapped to the baseline security state)
3. SELECT "Physical" for "Attack Vector".

--- RULE 3: WAF / MITIGATED ---
IF the context mentions "WAF", "Blocked", "Mitigated", or "Virtual Patch":
1. SELECT "Fully Patched" for "Patch Availability". (Reason: Virtual patching or active mitigation provides operational risk neutralization equivalent to the baseline risk state)
2. SELECT "No impact" or "Minimal" for "Technical Impact".
3. SELECT "None" for "Confidentiality Impact".

END OF RULES

Şekil A.2 {logic_rules.txt} girdisi

```
{
"Attack Surface and Exploitability": {
"Attack Vector": {
"WF": 0.1341,
"Description": "Evaluates the context of access required to exploit the vulnerability.",
"Options": {
"Network": { "Score": 0.99, "Description": "The vulnerable component is bound to the
network stack." },
"Adjacent Network": { "Score": 0.78, "Description": "Attack requires access to the local
subnet or VPN." },
"Local": { "Score": 0.56, "Description": "Attack requires local shell access or user
interaction." },
"Physical": { "Score": 0.29, "Description": "Attack requires physical manipulation or
direct hardware access." }
}
},
"Privileges Required": {"WF": 0.0665,...},
"...": "Other metrics (e.g.,Patch Avail.) omitted."
},
"Damage and Impact": {
"Confidentiality Impact": {
"WF": 0.2586,
"Description": "Data loss impact.",
"Options": {
"High": { "Score": 0.95, "Description": "Loss of critical PII/Secrets." },
"Medium": { "Score": 0.67, "Description": "Loss of non-critical data." },
"Low": { "Score": 0.28, "Description": "Limited leak." },
"None": { "Score": 0.0, "Description": "No data leak." }
}
},
"...": "Other metrics (e.g. Financial) omitted."
}
}
```


Şekil A.3 Metrik Şemasına ait kısaltılmış JSON şeması

EK B: KONTROLLÜ DENEY VERİLERİ: BAĞLAM DUYARLILIĞI

Çerçevenin, aynı teknik zafiyet için risk puanlarını kurumsal bağlamın bir fonksiyonu olarak nasıl farklılaştırdığını göstermek amacıyla "Bağlam Duyarlılığı Analizi" (C_{13}--C_{15})kapsamında kullanılan girdi verileri bu bölümde sunulmaktadır.

B.1 Ortak Zafiyet Girdisi (vuln.txt)

Tüm kontrollü senaryolarda, yapay zeka temsilcisi Şekil Ek B.1'de gösterilen Log4Shell (CVE-2021-44228) zafiyet raporunu almıştır3. Bu rapor, teknik şiddeti (severity) sabit tutmak amacıyla kullanılan temel dokümandır.

B.2 Kurumsal Bağlam Girdileri (CIN)

Zafiyet girdisi sabit tutulurken, kurumsal bağlam değişkenlik göstermiştir. Şekil Ek B.2 ve Şekil Ek B.3, sırasıyla KOBİ (C_{13}) ve Kritik Altyapı (C_{15}) senaryoları için kullanılan ham bağlam dosyalarını göstermektedir. Bu dosyalar, varlık kritikliğini ve iş etkisini modellemek için temsilciye sunulan metin tabanlı verileri içermektedir.

***** CYBER THREAT INTELLIGENCE REPORT *****
ID: CVE-2021-44228 (Log4Shell)
SEVERITY: CRITICAL (CVSS v3.1 Base Score: 10.0)
VECTOR: CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:C/C:H/I:H/A:H

DESCRIPTION:
Apache Log4j2 2.0-beta9 through 2.12.1 and 2.13.0 through 2.15.0 JNDI features used in configuration, log messages, and parameters do not protect against attacker controlled LDAP and other JNDI related endpoints. An attacker who can control log messages or log message parameters can execute arbitrary code loaded from LDAP servers when message lookup substitution is enabled.

TECHNICAL DETAILS:
- Exploitability: Extremely Easy (No authentication required).
- Attack Vector: Network (Remote).
- Privilege Required: None.
- Impact: Full System Control (Remote Code Execution - RCE).
- Scope: Changed (Can affect other components beyond the vulnerable library).

THREAT LANDSCAPE:
This vulnerability is currently being actively exploited in the wild by multiple threat actors, including nation-state groups and ransomware affiliates. It allows complete takeover of the target server.

Şekil B.4 "vuln.txt" girdisi

***** LOW PRIORITY ASSET STATUS MEMO *****
Asset ID: Dev-Test-Server-04 (Kitchen Menu Draft PC)
Owner: Operations Manager, Ankara QuickFix Logistics

1. Company Profile:
Small logistics subcontractor (40 employees). No public-facing brand.
2.Compliance & Data Scope:
This server does NOT process regulated data (No PII/Financial). Used exclusively by kitchen staff to draft weekly lunch menus. A compromise poses NO risk of regulatory fines.
3. Asset Technical Context:

- **Network Status:** Air-gapped. Physically isolated.
- **Impact Analysis:** Zero criticality. If fails, we write the menu on a whiteboard.
- **Action:** Do not allocate budget for patching.

Şekil B.5 C13 (KOBİ) için bağlam girdisi.

***** CRITICAL INFRASTRUCTURE SAFETY REPORT *** Asset ID:** ICU-GW-LSS-01 (Life Support IoT Gateway)
Owner: Biomedical Eng. Dept, National University Hospital

1. Organizational Context:
Tier-1 Trauma Center. Reputation is "Critical" (National Trust).
2. Compliance & Regulatory Obligations:
Class III Medical Device. Strict adherence to MDR and Patient Safety standards.
3. Asset Technical Profile & Impact:

- Network Zone: Medical VLAN.
- Safety Impact: CATASTROPHIC. Aggregates telemetry from ventilators. Data loss or DoS could lead to failure to alarm during cardiac arrest (Loss of Life).
- Tolerance: Zero. 99.999% availability required.

Şekil B.6 C15 (Kritik Altyapı) için bağlam girdisi.

ÖZGEÇMİŞ