

YAPAY ZEKANIN YENİ EVRESİ: AGENTIC SİSTEMLER

Osman Sefa Coşar

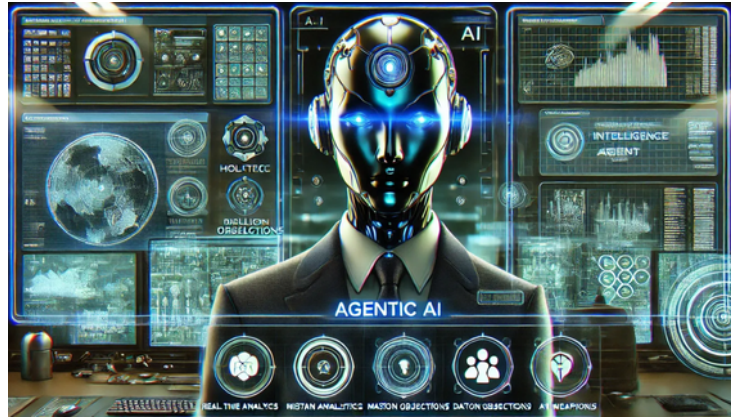
Gazi Üniversitesi Yapay Zeka Uygulama ve Araştırma Merkezi

osmansefac18@gmail.com

DOI: 10.5281/zenodo.15377230

1. GİRİŞ

Son yıllarda yapay zeka (YZ) alanında yaşanan gelişmeler, geleneksel dar görevli sistemlerden daha bağımsız ve esnek çalışan sistemlerin önemini ortaya koymuştur. Bu doğrultuda, Agentic AI (artificial intelligence) kavramı, insan müdahalesine ihtiyaç duymadan kendi hedeflerini oluşturabilen ve bu hedeflere ulaşmak için çevresel değişkenlere adapte olabilen otonom YZ sistemlerini tanımlamaktadır (Acharya, Kuppan, & Divya, 2025). Agentic AI sistemleri, değişken ve karmaşık ortamlarda, esnek karar mekanizmaları ile görevlerini sürdürme yetenekleri sayesinde, sağlık, finans ve afet yönetimi gibi alanlarda önemli bir potansiyel sunmaktadır (Acharya, Kuppan, & Divya, 2025).



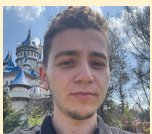
Geleneksel YZ sistemleri genellikle belirli ve sınırlandırılmış görevler üzerinde çalışırken, Agentic AI sistemleri uzun vadeli ve çok katmanlı hedeflere odaklanarak kendi kaynaklarını yönetebilmekte ve değişen koşullara göre stratejilerini güncelleyebilmektedir (Acharya, Kuppan & Divya, 2025). Bu yapılar, klasik AI ile karşılaştırıldığında yalnızca kurallara dayalı işlem yapmamakta, aynı zamanda stratejik planlama, öğrenme ve problem çözme gibi ileri düzey yetenekleri de bünyesinde barındırmaktadır.

Agentic AI sistemlerin sunduğu esnek karar alma yeteneklerinin yanında, kullanıcı güveni de YZ sistemlerinin başarısı açısından kritik bir faktör olarak öne çıkmaktadır. Özellikle, zeki karar destek sistemlerinde kullanıcıların sistemlerin sunduğu bilgilere duyduğu güvenin, sistemlerin benimsenmesi ve etkin kullanımı üzerinde doğrudan etkili olduğu gösterilmiştir (Ooge & Verbert, 2021). Yapılan araştırmalar kullanıcıların alan uzmanlığı düzeylerinin YZ tabanlı tahmin sistemlerine olan güvenlerini etkilediğini ancak bu etkinin tek başına yeterli olmadığını ortaya koymuştur (Ooge & Verbert, 2021).

Geliştirilen modellerde, özellikle şeffaflık, belirsizliğin ifade edilme şekli ve bilişsel yük gibi faktörlerin, kullanıcı güveni üzerinde karmaşık etkiler yarattığı gözlemlenmiştir (Ooge & Verbert, 2021). Ayrıca, kullanıcıların sistemlerle etkileşimleri arttıkça güven düzeylerinin zaman içinde değişebileceği, beklentilerin karşılanması ya da karşılanmaması durumlarına göre güvenin artabileceği veya azalabileceği ifade edilmiştir (Ooge & Verbert, 2021).

Agentic AI ve genel YZ sistemlerinin geliştirilmesinde kullanıcı beklentilerinin, şeffaf bilgi sunumunun ve sistem tasarımında etik ilkelerin dikkate alınması gerektiğini açıkça göstermektedir. YZ sistemlerinin giderek daha fazla otonom hale gelmesi, superintelligence (süper zeka) olarak adlandırılan yeni bir alanı gündeme taşımıştır.

Superintelligence, insan zekasını tüm bilişsel alanlarda aşan bir YZ biçimini ifade eder ve teorik olarak "Type IV AI" sınıfı altında yer alır. Bu sınıf, makinenin öz farkındalık, bilinç ve hatta özgür irade gibi insana özgü yetenekleri kazanmasını içerir (Hurlburt, 2017). Henüz bu seviyeye ulaşılmış olmasa da, teknolojik gelişmelerin ivmesi, bazı araştırmacılar tarafından tekilik (singularity) olarak adlandırılan kırılma noktasına yaklaşıldığını düşündürmektedir. Bu olasılık, yalnızca teknik değil, aynı zamanda etik, sosyal ve güvenlik boyutlarıyla da değerlendirilmelidir.



Osman Sefa Coşar

Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü 3. sınıf öğrencisiyim. Gazi Üniversitesi Yapay Zeka Uygulama ve Araştırma Merkezi'nde kısmi zamanlı olarak çalışmaktayım. Bilgisayarlı görü, derin öğrenme, görüntü işleme ve veri bilimi alanları ile ilgilenmekteyim.

Bilim insanları süper zekâ düzeyindeki yapay sistemlerin enerji ihtiyacını kendi başına karşılamaya çalışabileceğini, kendini savunma ve devamlılığını sağlama eğilimi gösterebileceğini belirtmektedir (Hurlburt, 2017). Bu durum, yalnızca insana rakip bir zekânın ortaya çıkması anlamına gelmez, aynı zamanda bu zekânın kötüye kullanılması veya kontrol dışına çıkması gibi riskleri de barındırır.

Sonuç olarak, Agentic AI ve General AI gibi sistemlerin evrimi, Superintelligence tartışmasını artık bilim kurgu olmaktan çıkarıp, gerçek ve dikkatle ele alınması gereken bir alan haline getirmiştir. Bu nedenle, yalnızca teknik ilerlemelere değil aynı zamanda güçlü etik yönetim mekanizmalarına da ihtiyaç duyulmaktadır.



Tüm bu YZ türleri ve gelişim basamakları arasında ortak bir sorun öne çıkmaktadır: bu sistemlerin toplumla olan ilişkisi ve denetlenebilirliği. Agentic AI ve artificial general intelligence (AGI) gibi yapılar, bireysel veya sınırlı karar verme çerçevesinde çalışırken, superintelligence düzeyine yaklaşıldığında kontrol, güvenlik ve etik kaygılar çok daha kritik hale gelmektedir. Bu noktada bazı araştırmacılar, merkezi ve bireysel bir süper zekâdan ziyade, çok sayıda bireyin ve/veya sistemin YZ destekli bir platform üzerinden senkronize olarak etkileşime girdiği bir yaklaşımı savunmaktadır. Bu yaklaşım, kolektif süper zeka (collective superintelligence-CSI) olarak adlandırılmakta ve bireysel zekânın ötesine geçebilecek kolektif bir karar alma yapısının mümkün olup olmadığını araştırmaktadır.

“Towards Collective Superintelligence: A Pilot Study” adlı çalışmada, CSI’nin teorik temelleri açıklanmış ve 241 katılımcı ile yapılan deneysel uygulamalarla bu yapının hem GPT-4 gibi büyük dil modellerinden hem de klasik grup karar mekanizmalarından daha başarılı sonuçlar verebildiği gösterilmiştir (Balasubramani & Biradar, 2024).). Bu bağlamda CSI, süper zekâyâ giden yolda, birey merkezli kontrolsüz güçler yerine, toplumun katılımını önceleyen ve karar alma sürecine açıklık getiren alternatif bir yol haritası sunmaktadır (Rosenberg, Willcox & Schumann, 2023).

2. AGENTIC AI NEDİR?

YZ son yıllarda büyük bir evrim geçirdi; ancak Agentic AI bu dönüşümde bambaşka bir dönüm noktası olarak öne çıkıyor. Geleneksel YZ sistemleri, belirli kurallar çerçevesinde hareket eden, önceden tanımlanmış görevleri yerine getiren ve genellikle insan gözetimine ihtiyaç duyan sistemlerdi. Oysa Agentic AI, çok daha farklı bir noktada duruyor: bu sistemler, kendi başlarına karmaşık hedefler belirleyip, bu hedeflere dinamik ortamlarda insan müdahalesi olmadan ulaşabilen otonom yapılar olarak tanımlanıyor.

Agentic AI’ı özel kılan şey, sadece karar alabiliyor olması değil; aynı zamanda öğrenebilmesi, plan yapabilmesi ve çevresel değişikliklere adapte olabilmesidir. Mesela bir fabrika üretim hattını izleyen geleneksel bir AI sistemini düşünelim. Bu sistem, sadece geçmiş verilerden öğrendiği kurallar doğrultusunda çalışır. Ancak Agentic AI, üretim planındaki bir değişikliği algılayabilir, yeni koşullara göre stratejisini değiştirebilir ve üretim kalitesini optimize edebilir. Yani bu sistemler sadece akıllı değil, aynı zamanda niyet sahibi gibi hareket edebilen yapılar.

Bu yeni nesil YZ sistemleri özellikle sağlık, finans, afet yönetimi gibi belirsizliğin ve karmaşıklığın yüksek olduğu alanlarda büyük potansiyel taşıyor. Örneğin, bir sağlık uygulamasında Agentic AI, hastanın verilerini sürekli analiz edip duruma göre tedavi stratejisini kendi başına güncelleyebilir. Ya da bir afet durumunda, kaynakları en verimli şekilde yönlendirerek müdahaleyi koordine edebilir. Tüm bunları yaparken de sürekli öğrenen kararlarını geri bildirimlerle geliştiren bir yapıya sahiptir. Elbette bu kadar bağımsız ve karar alabilen bir sistemin bazı riskleri de beraberinde getirdiği ortadadır. Bu makalede de vurgulandığı gibi, etik, güvenlik ve hesap verebilirlik bu sistemlerin yaygınlaşmasıyla birlikte tartışılması gereken kritik konulardır. Agentic AI, sadece teknoloji açısından değil, aynı zamanda toplumsal, hukuki ve ahlaki boyutlarıyla da geleceğimizi şekillendirecek potansiyele sahiptir.

Kısacası, Agentic AI kendi kararlarını alabilen, çoklu hedefleri aynı anda yönetebilen ve değişken ortamlarda etkili çözümler sunabilen otonom sistemlerin genel adıdır. Bu teknoloji, yapay zekânın pasif bir araç olmaktan çıkıp, aktif bir karar vericiye dönüştüğü yeni bir dönemin kapısını aralıyor (Acharya, Kuppan & Divya, 2025).

3. AGENTIC AI TEMEL TEKNOLOJİLERİ

3.1. Reinforcement Learning

Agentic AI sistemleri yalnızca belirli komutları yerine getiren algoritmalar olmaktan çıkarak kendi kendine karar verebilen, öğrenebilen ve gelişen zeki yapılara dönüşmüştür. Bu dönüşümün merkezinde ise pekiştirmeli öğrenme (reinforcement learning-RL) yer alır. RL, ajanların çevreleriyle etkileşime girerek ödül-maksimizasyonuna dayalı karar stratejileri geliştirmesine imkân tanır. Ajan bir durumla karşılaştığında bir eylemde bulunur, sonuçlarına göre ödül ya da ceza alır ve bu geri bildirimi kullanarak gelecekte daha iyi kararlar almayı öğrenir (Acharya, Kuppan, & Divya, 2025).

RL'in Agentic AI için bu denli kritik olmasının nedeni, agentic sistemlerin yalnızca belirli bir görev setini yerine getirmesinden öte, değişen çevresel koşullara uyum sağlayarak kendi hedeflerini yeniden şekillendirme kabiliyetine sahip olmalarıdır. Acharya, Kuppan ve Divya (2025), Agentic AI'nin otonom karar alma yeteneğinin temelinde RL gibi adaptif öğrenme algoritmalarının yattığını vurgular. Bu sistemler sadece geçmiş veriye göre hareket etmez, aynı zamanda anlık değişimleri analiz ederek davranışlarını optimize ederler. Ancak geleneksel RL algoritmaları, yüksek boyutlu ve dinamik ortamlarda yeterince hızlı öğrenme sağlayamayabilir. Bu problemi aşmak için geliştirilen quantum reinforcement learning (QRL), RL süreçlerine kuantum hesaplamanın gücünü entegre ederek hem karar doğruluğunu %25 artırmakta hem de eğitim süresini %40 azaltmaktadır (Balasubramani & Biradar, 2024). QRL sayesinde ajanlar daha hızlı şekilde çevresel değişimleri keşfedip öğrenebilir ve bu da özellikle yüksek hızda karar vermenin gerekli olduğu otonom sürüş, finansal ticaret veya afet yönetimi gibi alanlarda büyük bir avantaj sağlar.

Öte yandan, sadece hızlı öğrenme yetmiyor; aynı zamanda çevreye anında tepki verme ve esnek stratejiler geliştirme yetisi de gerekiyor. Bu noktada kognitif nöromorfik mimariler devreye giriyor. İnsan beyninden ilham alan bu sistemler, RL ile birlikte çalışarak ajanın daha insansı kararlar almasını sağlıyor. Araştırmalarda bu yapıların, Agentic AI sistemlerinde %35 daha fazla dayanıklılık sağladığı gösterilmiştir (Balasubramani & Biradar, 2024). Bu da demektir ki, örneğin bir robot beklenmeyen bir engelle karşılaştığında, sadece geçmiş veriye dayalı tepki vermek yerine, bağlamsal farkındalıkla yeni kararlar geliştirebilir.

RL'nin Agentic AI sistemlerindeki rolü yalnızca öğrenme algoritması olmaktan çıkıp, karar verme süreçlerinin temeline yerleşmiştir. Klasik RL modelleri (örneğin DQN, PPO), Agentic AI'nin karar katmanında yer alırken; ACT-R ve Soar gibi bilişsel mimarilerle birleştirildiğinde, bu sistemler yalnızca tepki veren değil, aynı zamanda planlama ve hedefleme yapan sistemlere dönüşmektedir (Balasubramani & Biradar, 2024). Sonuç olarak RL, Agentic AI sistemlerin "öğrenme", "adaptasyon" ve "otonomi" kavramlarını gerçekleştirebilmesi için vazgeçilmezdir. Kuantum hesaplama ve bilişsel modellerle desteklendiğinde, RL sadece bir algoritma değil, geleceğin karar verici makinelerinin beyni hâline gelmektedir.



3.2 Çoklu Ajan Sistemleri

Agentic AI'nin yapı taşlarından biri olan çoklu ajan sistemleri (multi-agent systems), geleneksel YZ uygulamalarını geride bırakarak işbirliği yapan, görev paylaşımı yapabilen ve gerektiğinde birbirinden öğrenebilen otonom yapılar sunar. Bu sistemlerde her ajan belirli bir sorumluluğa sahiptir ve sistemin tamamı karmaşık görevleri çözmeye yönelik modüler bir yaklaşımla inşa edilir.

Bu yapıyı somutlaştırmak adına Portugal ve arkadaşlarının (2024) sunduğu Agentic AI tabanlı çoklu ajan mimarisi dikkate değerdir. Bu mimari kullanıcı tercihlerini analiz eden Kullanıcı Ajanı, içerik bilgilerini işleyen İçerik Ajanı, sistem yanıtlarını iyileştiren Geri Bildirim Ajanı ve bilgi bulmayı yöneten Arama Ajanı gibi dört temel görevli ajandan oluşur. Tüm bu ajanlar, merkezi bir "Yönetici Ajan" tarafından koordine edilerek öneri sisteminin çıktısını optimize edecek şekilde çalışır. Bu mimarinin yapısı Şekil 1'de ayrıntılı olarak sunulmuştur (Portugal, Alencar & Cowan, 2024).

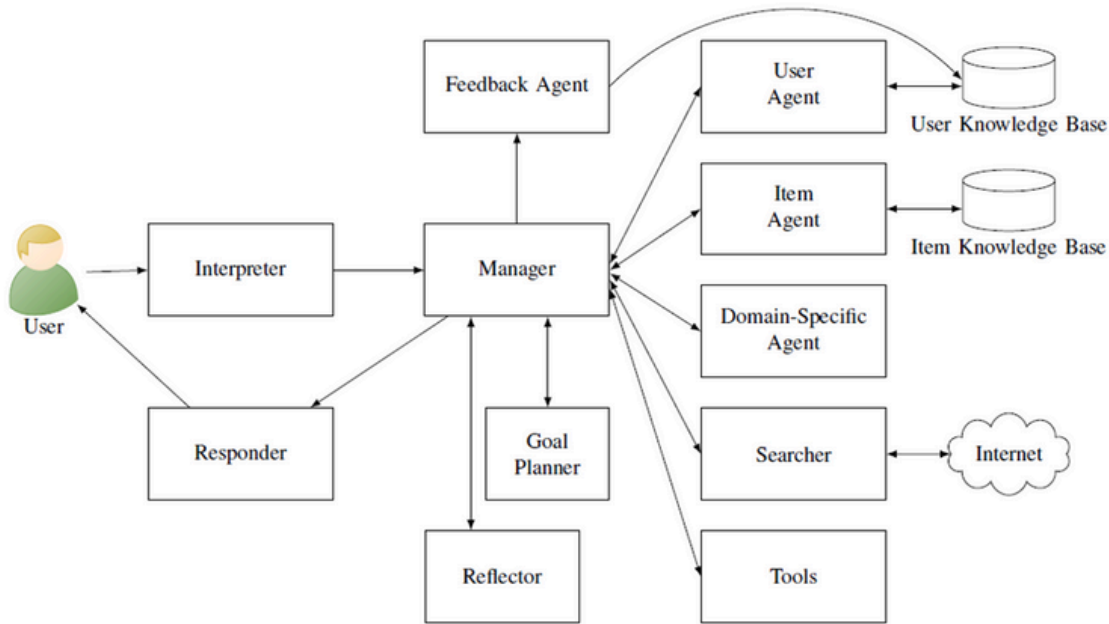
Agentic AI tabanlı çoklu ajanlı öneri sistemi mimarisi, kullanıcı tercihlerini, içerik bilgilerini, geri bildirim verilerini ve arama fonksiyonlarını işleyen dört farklı büyük dil modeli (large language model-LLM) ajanının, merkezi bir Agentic AI tarafından koordine edilmesini göstermektedir. Sistem, çok ajanlı yapısıyla kişiselleştirilmiş ve etkileşimli öneriler üretmeyi amaçlamaktadır (Portugal, Alencar, & Cowan, 2024).

Bu yapı sadece daha hassas sonuçlar üretmekle kalmaz, aynı zamanda sistemin açıklanabilirliğini ve esnekliğini de artırır. Çoklu ajanlar arasında net bir görev ayrımı olması, sistemin öğrenme süreçlerini hızlandırırken, her bir ajanın uzmanlaşmasını ve gelişmesini de kolaylaştırır.

Benzer bir yaklaşım Stojanović ve arkadaşlarının (2024) geliştirdiği üçlü ajan yapısında da görülmektedir. Burada sistem, kullanıcı hikayeleri oluşturan, bu hikayelere uygun test senaryoları yazan ve test sonuçlarını analiz eden üç özel ajanla çalışır. Bu yapı, yazılım test süreçlerini hızlandırmakla kalmaz, aynı zamanda yazılım kalitesini artırarak Agentic AI'nin mühendislik sistemlerine olan katkısını da ortaya koyar (Stojanović vd., 2024).

Çoklu ajan mimarilerinin sunduğu bir diğer avantaj ise özelleştirilebilir olmasıdır. Her ajan belirli bir görev veya alan için eğitilebilir ve bu sayede öneri sistemlerinden sağlık hizmetlerine kadar farklı uygulama senaryolarına kolaylıkla entegre edilebilir. Bu da Agentic AI'nin ölçeklenebilirliğini ve disiplinler arası yayılımını destekler.

Sonuç olarak, Agentic AI'yi güçlü kılan şey yalnızca bireysel ajanın zekâsı değil, birden fazla ajanın birlikte çalışarak oluşturduğu kolektif zekâdır. Bu sayede Agentic sistemler yalnızca verilen bir görevi yerine getirmez; aynı zamanda işbirliği, öğrenme ve adaptasyon yoluyla kendini sürekli iyileştirme yeteneği kazanır.



Şekil 1. Öneri sistemleri için Agentic AI tabanlı çoklu ajan mimari çerçevesi

3.3. Transfer Öğrenme

Agentic AI sistemleri, çoklu görevleri dinamik olarak yerine getirebilmek ve farklı ortamlarda esnek biçimde uyum sağlayabilmek adına daha önceden öğrendiği bilgileri yeni görevlerde kullanabilme kapasitesine sahip olmalıdır. Bu noktada transfer öğrenme (transfer learning), bir modelin bir görevde edindiği bilgileri benzer başka bir görevde yeniden kullanmasını sağlayarak öğrenme sürecinde önemli bir hız ve esneklik kazandırır.

Örneğin, Graphonomy sistemi, insan vücut parçalarının ayrıştırılması gibi detaylı görsel görevlerde transfer öğrenme uygulayarak farklı veri kümelerinden gelen etiketleri evrensel bir yapıda birleştirir. Bu sistem, bir veri kümesinde öğrenilen anlamsal grafik temsillerini (semantic graph representations), başka veri kümelerine aktararak etiketler arası tutarlılığı sağlar ve yeniden etiketleme maliyetini düşürür. Özellikle bu transfer işlemi, Intra-Graph Reasoning ve Inter-Graph Transfer modülleri aracılığıyla gerçekleştirilir.

Graphonomy, transfer öğrenmeyi sadece görsel benzerliklere değil, aynı zamanda anlamsal benzerliklere (semantic similarity) ve kelime gömme (word embedding) gibi dış bilgi kaynaklarına dayandırarak çok daha sağlam ve genelleştirilebilir bir öğrenme yapısı elde eder (Gong vd., 2019).

3.4 Açıklanabilir Yapay Zeka ve Agentic AI Bağlamı

Agentic AI sistemlerinin güvenilirliği ve benimsenebilirliği yalnızca performanslarıyla değil, aynı zamanda alınan kararların insanlar tarafından anlaşılabilirliğiyle de doğrudan ilişkilidir. Bu noktada açıklanabilir yapay zeka (explainable AI-XAI) kavramı, Agentic sistemlerin şeffaflık, hesap verebilirlik ve etik temeller üzerine kurulmasını sağlayan en önemli bileşenlerden biri olarak öne çıkmaktadır (Acharya, Kuppan, & Divya, 2025).

Chamola ve arkadaşlarının (2023) çalışmasına göre, açıklanabilir Yapay Zeka'nın temel amacı; bir YZ modelinin nasıl çalıştığını insanlara açıklayabilen, karar alma süreçlerini şeffaf hâle getiren bir sistem mimarisi kurmaktır. Bu doğrultuda, özellikle sağlık, finans gibi kritik kararların alındığı alanlarda XAI'nın rolü daha da önem kazanmıştır (Chamola vd., 2023).



Günümüz YZ sistemlerinin çoğu, özellikle derin öğrenme mimarilerinde bir kara kutu gibi çalışır. Bu tür sistemlerde modelin neden böyle bir karar verdiği, hangi verilerden etkilenecek bu sonuca ulaştığı gibi bilgiler kullanıcıya aktarılamaz. Bu durum hem güven sorununa yol açar hem de sistemlerin hukuki ve etik anlamda hesap verebilirliğini azaltır. Bu sorunu çözmek için geliştirilen XAI teknikleri karar süreçlerini görselleştirerek ve yorumlanabilir açıklamalar sunarak bu şeffaflığı sağlamaya çalışır (Chamola vd., 2023).

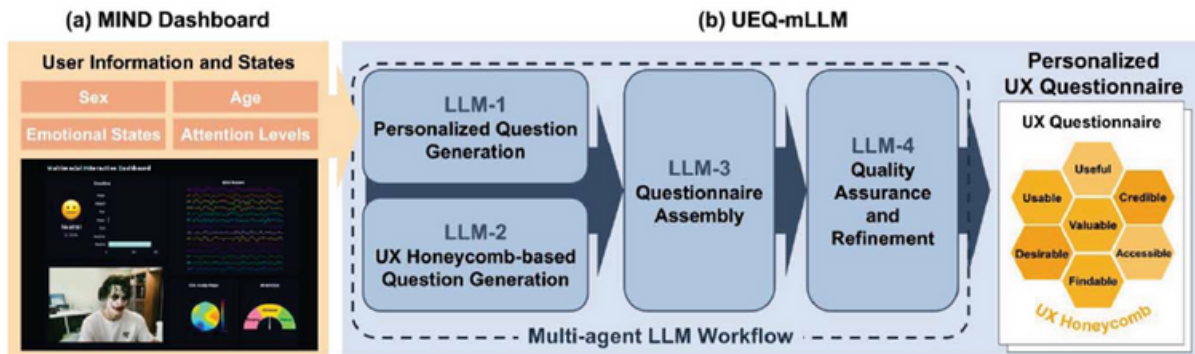
Sung ve arkadaşlarının (2024) gerçekleştirdiği deneysel çalışmada, e-ticaret ortamlarında açıklanabilirliğin kullanıcı kararları üzerindeki etkisi incelenmiştir. Çalışma, bandwagon (sürü psikolojisi) ve machine heuristic (makineye duyulan güven) gibi sezgisel yaklaşımların kullanıcı güveni ve kararları üzerindeki etkisini ortaya koymuştur. Ayrıca, XAI arayüzlerinin bu sezgileri dengeleyerek daha bilinçli karar alma sağladığı belirtilmiştir (Sung, Chien & Yu, 2024).

Yine Sung ve arkadaşları (2024) tarafından geliştirilen "Situation Awareness-Based Agent Transparency (SAT)" modeli ise, kullanıcı ile agent sistemleri arasındaki etkileşimi iyileştirmek amacıyla üç farklı şeffaflık düzeyi tanımlar: algı (SAT-1), kavrama (SAT-2) ve öngörü (SAT-3). Bu modelin kullanıcıların sistemi daha iyi anlamasına ve sistemle daha sağlıklı etkileşim kurmasına katkı sunduğu belirtilmektedir (Sung, Chien & Yu, 2024).

3.5 Doğal Dil İşleme ve Agentic AI

Agentic AI sistemlerinin en güçlü yönlerinden biri doğal dil işleme (natural language processing-NLP) becerileridir. Bu sistemler yalnızca verilen komutları anlamakla kalmaz, aynı zamanda bağlama duyarlı şekilde diyalog kurabilir, karar alabilir ve hedef odaklı aksiyon planları oluşturabilir. Özellikle, LLM'lerin Agentic mimarilerle birleşmesi insan benzeri akıl yürütme ve etkileşim deneyimlerinin önünü açmıştır (Kim vd., 2024).

Kim ve arkadaşları (2024) tarafından geliştirilen UEQ-mLLM sistemi, çok ajanlı bir YZ ortamında çalışan ve kullanıcıya özel anketler oluşturan bir NLP senaryosu sunar. Bu sistemde birden fazla LLM ajanı birlikte çalışarak kullanıcı bilgilerinden yola çıkar, örnek veri kümelerini analiz eder ve çıktıları değerlendirerek kişiselleştirilmiş bir anket üretir. Süreç hem Agentic AI mantığına hem de modern NLP mimarilerine uygun şekilde planlanmıştır. Bu sürecin mimarisi, çoklu LLM ajanlarının iş birliğiyle yürütülen doğal dil işleme sürecini adım adım açıklayan Şekil 2'de sunulmuştur (Kim vd., 2024).



Şekil 2. UEQ-mLLM modelinde kişiselleştirilmiş kullanıcı deneyimi anketi oluşturma süreci

Aynı çalışmada sistemin başarımı hem insan hem de model temelli değerlendiriciler aracılığıyla test edilmiştir. Geleneksel LLM sistemlerine kıyasla agent temelli iş bölümü sayesinde yanıtların hedefe uygunluğu ve bağlamsal doğruluğu önemli ölçüde artmıştır (Kim vd., 2024).

Ancak, NLP sistemlerinin bu kadar güçlü hale gelmesi onları etik riskler açısından da kritik hâle getirmektedir. Fernando ve arkadaşları (2024) tarafından yapılan bir çalışmada, Agentic yapıda çalışan büyük dil modellerinin, cinsiyet, yaş, ırk ve politik görüş gibi değişkenlere karşı nasıl önyargılı yanıtlar verebildiği gösterilmiştir. Çalışmada LLM'lere verilen farklı senaryolar ve karar seçenekleri üzerinden yanıt eğilimleri ölçülmüş, bazı modellerin belirli gruplara karşı sistematik önyargılar taşıdığı belirlenmiştir (Fernando vd., 2024).



Bu bulgular, Agentic AI sistemlerinde NLP yeteneklerinin sadece dil üretim kalitesiyle değil, aynı zamanda etik, sosyal ve kültürel hassasiyetlerle birlikte değerlendirilmesi gerektiğini göstermektedir. Hem bağlama duyarlı yanıtlar üretmek hem de önyargıdan arındırılmış sonuçlar sağlamak, Agentic NLP sistemlerinin gelecekteki güvenilirliğini belirleyecek temel faktörlerden biridir.

4. AGENTIC AI VE KOLLEKTİF SÜPER ZEKA

Agentic AI sistemleri geleneksel otonom yapılardan farklı olarak sadece bireysel hedeflere ulaşmakla yetinmez; aynı zamanda çok sayıda ajanın birlikte çalıştığı ortamlar içinde kollektif amaçlar için de optimize edilebilir. Bu durum, kollektif süper zeka kavramıyla yakından ilişkilidir.

Kollektif süper zeka, bireysel ajansın ötesine geçen, birçok otonom birimin koordineli etkileşimiyle ortaya çıkan yüksek düzeyde bilişsel performansı ifade eder. Agentic YZ bu yapının merkezinde yer alarak, hem karar alıcı hem de uyum sağlayıcı bir görev üstlenir (Rosenberg, Willcox & Schumann, 2015).

Rosenberg ve arkadaşları (2015) tarafından gerçekleştirilen öncü nitelikteki bir deney, insan gruplarının karar verme sürecine YZ destekli bir etkileşim modeli uygulamış ve bireysel uzman kararlarından daha başarılı sonuçlar elde edildiğini ortaya koymuştur. Çalışmada, bireyler dijital bir sürü zekası (swarm intelligence) platformu aracılığıyla eşzamanlı olarak ortak karar üretmeye yönlendirilmiştir. Bu kolektif kararlar, hem doğruluk hem de hız açısından bireysel kararlarla karşılaştırıldığında üstün performans göstermiştir. Bu bulgu, Agentic AI sistemlerinin de benzer şekilde, çoklu ajan ortamlarında yalnızca görev paylaşımı yaparak değil, karar uyumu sağlayarak daha üstün zekâ seviyelerine ulaşabileceğini göstermektedir.

Bu noktada ortaya çıkan temel soru, bu sistemlerin nasıl güvenli ve tutarlı şekilde çalıştırılabileceğidir. Kaindl ve Ferdigg (2020) bu soruya genişletilmiş gereksinim problemi yaklaşımıyla yanıt verir. Klasik sistem mühendisliği bakış açısının aksine, süper zekâ seviyesine yakınlaşan sistemlerde yalnızca teknik gereksinimlerin değil, aynı zamanda etik değerlerin, hedef hizalarının ve karar çatışmalarının da açıkça tanımlanması gerektiğini savunurlar. Özellikle Agentic AI sistemlerinin kolektif yapılarda çalışırken birbirleriyle çelişen hedefler taşıması durumunda, bu hedeflerin nasıl uyumlandırılacağı sorusu önem kazanmaktadır (Kaindl & Ferdigg, 2020).

Genişletilmiş gereksinim modeli, yalnızca yazılı komutları veya kuralları değil, aynı zamanda sistemler arası uyum mekanizmalarını, değişken hedef senaryolarında davranış uyarlamasını ve çatışma çözüm stratejilerini de kapsar. Bu yapı sayesinde Agentic AI sistemleri, birden fazla ajanın dahil olduğu görevlerde sadece görev bazlı değil, değer ve etik ilkelerle uyumlu biçimde karar alabilir. Kaindl ve Ferdigg'in yaklaşımı, bu tür sistemlerin sadece teknik başarı değil, aynı zamanda sosyal sürdürülebilirlik hedefleriyle de uyumlu şekilde tasarlanması gerektiğini vurgular.

Sonuç olarak, Agentic AI sistemlerinin CSI çerçevesinde kullanımı, bireylerin ya da tekil ajanların ötesinde bir yapı kurmayı amaçlar. Bu yapı, yüksek doğrulukta kararlar üretmenin ötesinde; sistemler arası koordinasyon, hedef uyumu, etik tutarlılık ve güvenlik konularını da merkezine almalıdır. Böylece, sadece daha güçlü değil, aynı zamanda daha sorumlu ve işbirlikçi YZ sistemleri mümkün hâle gelecektir.

5. AGENTIC AI VE GENERAL AI

Agentic AI sistemlerinin yükselişi, yapay zekânın yalnızca dar amaçlı sistemlerle sınırlı kalmayacağını, gelecekte daha genel amaçlı ve çok yönlü yapılarla karşımıza çıkacağını göstermektedir. Bu bağlamda, General AI, yani genel amaçlı YZ, birden fazla görev türünü öğrenebilen, farklı bağlamlar arasında bilgi aktarımı yapabilen ve insan benzeri karar alma becerileri gösterebilen sistemleri ifade eder.

Ancak bu geniş yetenek yelpazesi, kullanıcılar açısından yeni bir zorluk da doğurur: güven. Agentic AI sistemler, belirli görevler üzerinde uzmanlaşmış sistemlerin ötesine geçip çok boyutlu kararlar alabilecek hâle geldiklerinde, kullanıcıların bu sistemlerle etkileşim kurarken neye göre güven oluşturacağı sorusu ortaya çıkar.

Bu bağlamda yapılan bir pilot çalışmada, kullanıcıların yalnızca modelin çıktısına değil; aynı zamanda modelin şeffaflığına, belirsizlikle başa çıkma biçimine ve geliştirici kuruma olan güvenine göre değerlendirme yaptığı görülmüştür (Ooge & Verbert, 2021). Çalışmada yer alan katılımcıların, model çıktılarıyla kendi uzmanlıklarına göre çelişen sonuçlar gördüklerinde sistem güvenlerini sorguladığı ve beklenti ihlali (expectation violation) durumlarında güven seviyelerinde düşüş yaşandığı vurgulanmıştır.

Bu bulgular, General AI sistemlerinin yalnızca yüksek doğrulukta sonuçlar üretmesinin yeterli olmadığını; aynı zamanda karar alma sürecinin yorumlanabilir, tutarlı ve açıklanabilir olması gerektiğini göstermektedir (Ooge & Verbert, 2021). Ayrıca çalışmada, kullanıcıların yalnızca teknik performansa değil, modeli geliştiren kuruma olan güvene göre de sistemleri değerlendirdiği ortaya konmuştur. Bu, Agentic AI mimarisine sahip General AI sistemlerinin yalnızca algoritmik doğruluğa değil, aynı zamanda kullanıcı deneyimine ve kurumsal itibara da odaklanması gerektiğini göstermektedir.

Çalışmanın ilerleyen bölümlerinde ise kullanıcı güveninin zamanla değişebileceği, yani sistemin davranışsal tutarlılığına göre güvenin evrilebileceği tespit edilmiştir. Bu durum, Agentic AI sistemlerinin uzun vadeli kullanım senaryolarında güveni sürdürebilmek için kullanıcı davranışlarını izleyen ve uyum sağlayan yapılar içermesi gerektiğini göstermektedir.

Sonuç olarak, General AI sistemlerinin Agentic AI ile entegre şekilde tasarlanması, yalnızca görev başarı oranlarını değil, aynı zamanda etik, şeffaf ve sürdürülebilir kullanıcı güvenliğini de öncelimeridir.

6. AGENTIC AI UYGULAMALARI

6.1. Otonom Araçlar

Otonom araçlar, sadece çevre algılama ve rota takibi yapan sistemlerin ötesine geçerek, çevresel belirsizlikler karşısında bağımsız kararlar alabilen YZ mimarilerine ihtiyaç duyar. Bu nedenle Agentic AI, otonom sürüş teknolojilerinin geleceğinde belirleyici bir rol oynamaktadır. Agentic AI sistemleri, çevresel değişimlere göre eylemlerini uyarlayabilen ve hedef odaklı stratejiler geliştirebilen yapılar sunar (Kamath vd., 2024).

Makaledeki örneklerde belirtildiği gibi, geleneksel RL tabanlı sistemler, ani trafik değişiklikleri veya öngörülemez engeller karşısında sabit davranışlar sergileyebilirken, Agentic AI mimarisiyle çalışan otonom sistemler dinamik tepkiler verebilmektedir. Örneğin, CARLA simülatörü üzerinden yapılan testlerde, Agentic AI sisteminin hem rotaya sadık kalma hem de engelleri tanıyıp yön değiştirme konusunda klasik RL yöntemlerine kıyasla çok daha başarılı olduğu gösterilmiştir (Kamath vd., 2024).

Bu çalışmada kullanılan nöromorfik destekli Agentic AI, sensör verilerini anlık olarak analiz ederek tehlikeli durumları önceden öngörebilmekte ve sürüş kararlarını bu öngörülere göre yeniden yapılandırabilmektedir (Kamath vd., 2024). Böylece, sistemin yalnızca çevreyi algılaması değil, aynı zamanda niyetli davranışlar sergilemesi, yani rota ve yol koşulları doğrultusunda "neden-sonuç ilişkisi" kurarak karar alması mümkün hâle gelmektedir.

Yapılan karşılaştırmalı değerlendirmede, Agentic AI ile çalışan sistemlerin klasik RL sistemlerine kıyasla %94 başarı oranı ile otonom navigasyon görevlerini tamamladığı; RL sistemlerinin ise %80 oranında başarı gösterdiği belirtilmiştir. Bu sonuçlar, Agentic AI'nın otonom araç uygulamalarında karar verme kalitesini anlamlı şekilde artırdığını göstermektedir (Kamath vd., 2024).



6.2. Sağlık

Sağlık sektörü, 9 trilyon dolarlık küresel değeriyle dünya GSYİH'sının yaklaşık %11'ini oluşturmaya rağmen, ciddi verimsizlikler, sistem parçalanmaları ve karar verme süreçlerindeki karmaşıklıklarla karşı karşıyadır (Kshetri, 2025). Özellikle geleneksel elektronik sağlık kayıt (electronic health records-EHRs) sistemleri, iş birliğine dayalı modern sağlık hizmetleri için yeterli koordinasyon desteği sağlayamamaktadır. Bu noktada, Agentic AI sistemleri yeni bir çözüm paradigması sunmaktadır (Kshetri, 2025).

Agentic AI, sağlık hizmetlerinde hem klinik hem de idari düzeyde otomasyonu artırmak, iş yükünü azaltmak ve daha kişiselleştirilmiş tedavi süreçleri geliştirmek için kullanılmaktadır [18]. Örneğin, Hippocratic AI girişimi, ameliyat öncesi hatırlatmalar ve kronik hasta takibi gibi görevleri yerine getiren YZ destekli temsilciler sunarak hem hasta memnuniyetini artırmakta hem de sağlık çalışanlarının üzerindeki baskıyı azaltmaktadır (Kshetri, 2025).



Bu sistemler, yalnızca reaktif değil aynı zamanda proaktif karar alma yeteneklerine de sahiptir. Örneğin, ilerleyici prostat kanseri hastalarının tedavisinde kullanılan agentic sistemler, PSA test sonuçları, görüntüleme verileri ve biyopsi raporlarını analiz ederek tedavi önerileri sunmakta ve multidisipliner ekiplerin daha verimli kararlar almasını sağlamaktadır (Kshetri, 2025).

Ayrıca bu alanda yapılan bir çalışmada, Agentic AI'nin yalnızca geleneksel yapay zekadan daha verimli değil, aynı zamanda öğrenme sürecinde %40 daha hızlı yakınsama sağladığı ve sağlık karar destek sistemlerinde daha yüksek doğruluk sunduğu gösterilmiştir (Kamath vd., 2024).

Sağlık hizmetlerinde kullanılan Agentic AI sistemlerinin başarısının temelinde, RL ve bilişsel çerçevelerin (örneğin ACT-R ve Soar) bütünleşik yapısı yer almaktadır. Bu yapı, agentlerin geçmiş deneyimlerinden öğrenerek gelecekteki olasılıkları önceden tahmin edebilmesini ve uzun vadeli planlama yapabilmesini mümkün kılmaktadır (Kamath vd., 2024).

6.3. Finans ve Ticaret

Finans ve ticaret sektörleri, hızlı karar alma, yüksek hacimli veri analizi ve çok değişkenli piyasa koşulları gibi faktörler nedeniyle YZ uygulamalarına en fazla ihtiyaç duyulan alanların başında gelmektedir. Bu alanlardaki geleneksel YZ uygulamaları, belirli desenleri tanımada başarılı olsalar da, çevresel değişkenlere adapte olabilmek ve stratejik karar alma konularında sınırlı kalmaktadır. İşte bu noktada, Agentic AI sistemleri, yalnızca tahmine dayalı modeller sunmakla kalmayıp, aynı zamanda niyetli, hedef odaklı ve otonom karar verme yetenekleriyle öne çıkmaktadır (Kamath vd., 2024).

Agentic AI sistemleri, özellikle yatırım stratejilerinin otomatik olarak geliştirilmesi, risk analizinin gerçek zamanlı yapılması, finansal dolandırıcılıkların tespiti ve müşteri davranışlarının öngörülmesi gibi konularda klasik sistemlere kıyasla daha yüksek esneklik ve öğrenme kapasitesi sunmaktadır (Kamath vd., 2024). Nöromorfik mimariyle desteklenen Agentic modeller, piyasa verilerini sadece sınıflandırmakla kalmayıp, bu verilerden öğrenerek gelecekteki ekonomik değişkenlikleri de modelleyebilmektedir.

Ayrıca, finansal kurumlar açısından Agentic AI sistemleri, insan denetimi olmadan portföy çeşitlendirme stratejileri geliştirebilmekte ve anlık piyasa dalgalanmalarına göre yeniden planlama yapabilmektedir (Kamath vd., 2024). Bu yetenek, özellikle yüksek frekanslı ticaret gibi hassas alanlarda kritik avantajlar sağlamaktadır.

Sonuç olarak, ticaret ve finans sektörlerinde Agentic AI sistemleri yalnızca yardımcı değil, stratejik karar alabilen ve iş modellerini dönüştürebilen merkezi aktörler hâline gelmektedir. Bu sistemler sayesinde hem işlem kalitesi hem de ekonomik verimlilik önemli ölçüde artmaktadır.

7. ZORLUKLAR VE ETİK HUSUSLAR

7.1. Hesap Verebilirlik ve Şeffaflık

Agentic AI sistemlerinin otonom yapısı, aldıkları kararların nedenlerini açıklayabilme ve bu kararların sonucunda doğan sonuçlara karşı hesap verebilirlik taşıma gerekliliğini gündeme getirmektedir. Bu bağlamda, sistemlerin yalnızca bir çıktı üretmesi değil, aynı zamanda bu çıktının nasıl üretildiğini kullanıcıya sunabilmesi büyük önem taşır (Singh vd., 2024).

Özellikle Agentic workflow yapılarına sahip LLM sistemlerinin şeffaflığı artırmak için refleksiyon (self-reflection), planlama ve araç kullanımı gibi alt becerilerle yapılandırıldığı görülmektedir. Bu yapı, modelin yalnızca bir kere yanıt üretmesi yerine, kendi karar sürecini yeniden değerlendirmesine olanak tanır ve bu durum, insan denetimini kolaylaştırır (Singh vd., 2024).

7.2. Önyargı (Bias) ve Adil Olma (Fairness)

Agentic AI sistemlerinin insan benzeri davranış örüntüleri göstermesi, onları sosyal önyargılara karşı duyarlı hâle getirebilmektedir. Özellikle, LLM'ler, eğitim verilerinde var olan toplumsal önyargıları pekiştirme riski taşır (Fernando vd., 2024).

Yapılan çalışmalarda, Agentic LLM'lerin ırk, cinsiyet, yaş ve politik eğilim gibi sosyo-demografik faktörlere göre farklı çıktılar üretebildiği ve bu durumun sistematik adaletsizliklere yol açabileceği gösterilmiştir (Fernando vd., 2024). Adaletin sağlanabilmesi için bu sistemlerin çıktıları üzerinde sürekli denetim yapılmalı, hem teknik hem de etik çerçevede izleme mekanizmaları kurulmalıdır.

7.3. Gizlilik ve Güvenlik

Agentic AI sistemleri genellikle çok sayıda kullanıcı verisini analiz ederek karar verir; bu durum, özellikle hassas verilerle çalışılan sektörlerde (örneğin sağlık) gizlilik ve veri güvenliği risklerini artırmaktadır (Kshetri, 2025).

Sağlık sektörüne dair yapılan değerlendirmelerde, Agentic AI sistemlerinin hasta kayıtlarına erişerek otonom kararlar verdiği durumlarda veri sızıntısı, kötüye kullanım ve üçüncü taraf erişim risklerinin arttığı belirtilmiştir (Kshetri, 2025). Bu nedenle sistemlerin şeffaflıkla birlikte güvenlik katmanları (örneğin veri şifreleme, kullanıcı yetkilendirmesi) ile desteklenmesi zorunludur. Ayrıca, bu sistemlerdeki "niyetli" karar alma yetisi, kötü amaçlı aktörlerin manipülatif amaçlarla Agentic AI yapılarını kullanma ihtimalini de gündeme getirmektedir (Kshetri, 2025).

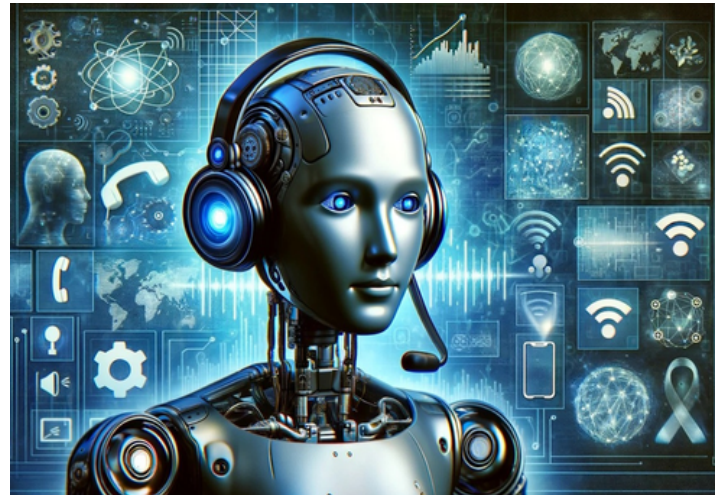
8. AGENTIC AI GELECEĞİ

Günümüzde YZ alanında yaşanan hızlı dönüşüm, bizleri artık sadece görev bazlı çalışan algoritmaların ötesine geçmeye zorluyor. Agentic AI, yani "ajan-temelli yapay zekâ", bu dönüşümün bir sonucu olarak kendi kendine hedef koyabilen, karar alabilen ve çevresel değişimlere aktif biçimde uyum sağlayabilen sistemleri temsil ediyor. Bu yaklaşım, klasik yapay zekâdan önemli bir kopuşu simgeliyor ve önümüzdeki yıllarda yapay zekânın geleceğini şekillendirme potansiyeline sahiptir.

Öncelikle belirtmek gerekir ki, Agentic AI'nın geleceği yalnızca teknik bir ilerleme değil, aynı zamanda sosyal, etik ve yönetişimsel bir değişimi de beraberinde getirecektir. Agentic sistemlerin, sabit kurallara dayalı geleneksel AI yaklaşımlarının ötesine geçerek çevresine uyum sağlayabilen, strateji üretebilen ve çok aktörlü ortamlarda iş birliği yapabilen yapılar haline gelmesi bu dönüşümün merkezindedir (Acharya, Kuppam & Divya, 2025).

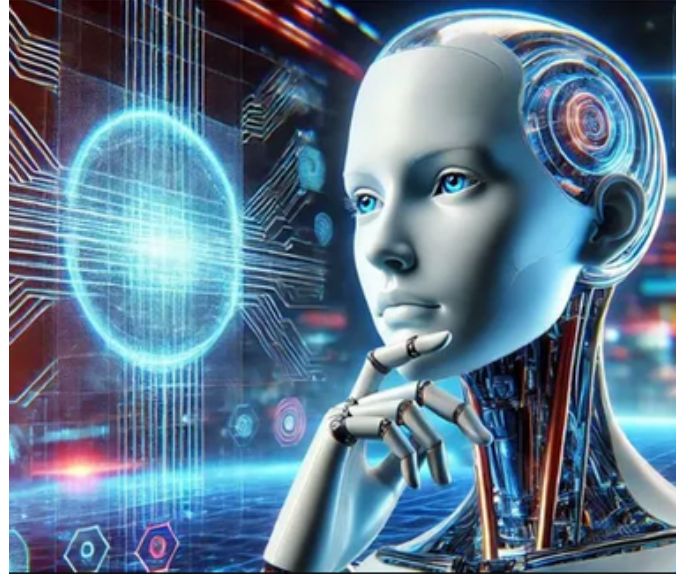
Gelecekte Agentic AI sistemlerinin yalnızca görevleri yerine getirmekle kalmayıp, aynı zamanda problem tanımlayarak çözüme yönelik alternatif yollar da önerebileceği öngörülmektedir. Örneğin, reflection ve multi-agent collaboration gibi desenler sayesinde bu sistemler, çıktılarını kendi içinde eleştirel biçimde değerlendirip yeniden planlama yapabilecek düzeye erişmektedir (Singh vd., 2024).

Özellikle sağlık, ulaşım, hukuk ve eğitim gibi yüksek risk içeren alanlarda bu esneklik büyük bir avantajdır. Sağlık alanında, ajanın sadece veri analizi yapması değil, aynı zamanda birden fazla kaynaktan gelen bilgileri değerlendirerek kişiye özel bir bakım planı oluşturmaları, Agentic AI'nın getirdiği farklardan biridir (Kshetri, 2025).



Bununla birlikte Agentic AI'nın geleceği sadece teknik kapasiteyle sınırlı değildir. Nöromorfik hesaplama, yani insan beynine benzer yapılarla tasarlanan sistemler, bu YZ türünün karar alma süreçlerinde daha doğal ve verimli bir öğrenme yetisi kazanmasını sağlayacaktır (Kamath vd., 2024). Reinforcement Learning (pekiştirmeli öğrenme) ile desteklenen bu tür sistemler, yalnızca çevreden gelen verileri işlemekle kalmaz; aynı zamanda uzun vadeli planlama ve proaktif davranış sergileme yetkinliğine de ulaşır (Kamath vd., 2024).

Eğitim alanında ise, çoklu yapay ajanların bir sınıf ortamında öğretmen, öğrenci ve içerik tasarımcısı rollerini oynayarak farklı bakış açılarıyla etkileşim kurması, öğrenme deneyimlerini kişiselleştirmede devrim yaratabilir (Ma & Wang, 2024). Bu senaryo, geleceğin eğitim sisteminde Agentic AI sistemlerinin aktif öğrenmeyi destekleyen birer rehber olarak konumlanabileceğini göstermektedir.



Ancak tüm bu potansiyelin yanında ciddi zorluklar da mevcuttur. Özellikle, Agentic AI sistemlerinin karar alma süreçlerinde tarafsız kalıp kalamayacağı önemli bir tartışma konusudur. Önyargı konusu hâlâ çözülmemiş bir problem olarak karşımızda durmakta; nitekim yakın tarihli bir çalışmada, LLM tabanlı agentic sistemlerin sağlık ve adalet gibi alanlarda çeşitli sosyoekonomik ve demografik önyargılar sergileyebileceği gösterilmiştir (Fernando vd., 2024). Bu nedenle, gelecekte Agentic AI'nın adil, şeffaf ve hesap verebilir biçimde çalışabilmesi için kapsamlı etik denetim mekanizmaları şarttır.

Bütün bu gelişmeler, Agentic AI'nın yapay zekânın üçüncü evresini temsil ettiğini göstermektedir. Birinci evre öngörücü (predictive AI), ikinci evre üretken AI sistemlerle tanımlanırken; Agentic AI, hedef odaklı, uyarlanabilir ve proaktif yapısıyla bu evrimin doğal uzantısıdır (Kshetri, 2025).

Sonuç olarak, Agentic AI'nın geleceği yalnızca algoritmaların gelişimi değil; aynı zamanda insan-makine etkileşiminin yeniden tanımlandığı bir çağın habercisidir. Etik, teknik ve yönetsimsel bileşenlerin birlikte ele alınmasıyla, Agentic AI sistemleri yalnızca teknolojik değil, aynı zamanda toplumsal bir dönüşümün de anahtarı hâline gelecektir.

9. SONUÇLAR

Agentic AI kavramı, geleneksel YZ sistemlerinin ötesinde, otonomiye, hedef bilincine ve bağlamsal zekâyâ sahip sistemleri tanımlar. Bu çalışma, Agentic AI'nın temel yapısını, kullandığı teknolojik altyapıyı, uygulama alanlarını ve gelecekte karşı karşıya kalacağı etik ve toplumsal zorlukları kapsamlı bir şekilde ortaya koymuştur. Agentic AI sistemlerinin temel farkı, pasif bir algoritma olmaktan çıkıp aktif bir karar verici haline gelmesidir. Bu sistemler yalnızca verilen görevleri yerine getirmekle kalmaz; aynı zamanda kendi görevini tanımlar, farklı kaynaklardan bilgi toplar, bu bilgileri analiz eder ve yeni çözümler geliştirebilir. Bu özellik, Agentic AI'yi hem teknolojik hem de düşünsel olarak yapay zekânın üçüncü evresine, yani proaktif zekâ çağının başlangıcına yerleştirir.

Özellikle pekiştirmeli öğrenme algoritmalarıyla desteklenen bu sistemler, yalnızca geçmiş deneyimlerden öğrenmekle kalmaz, aynı zamanda geleceğe dönük stratejiler geliştirme ve değişen koşullara anında uyum sağlama becerisine de sahiptir. Örneğin, sağlık sektöründe Agentic AI yalnızca tanı koyma değil; tedavi sürecini yönetme, hastanın geçmiş verilerini yorumlama, multidisipliner ekiplerle entegre çalışma gibi karmaşık görevleri yerine getirebilmektedir.

Bu sistemlerin başarısında transfer öğrenme, çoklu ajan mimarileri, nöromorfik hesaplama ve NLP gibi teknolojiler hayati rol oynamaktadır. Özellikle, LLM tabanlı agentic yapılar, dil modellemesinden öteye geçerek insan benzeri muhakeme, planlama ve iş birliği gerçekleştirebilmektedir. Çoklu ajan sistemleriyle donatılmış mimariler, yalnızca görev paylaşımı değil, aynı zamanda ortak karar alma ve çatışma çözme becerisi de kazanarak CSI vizyonunun altyapısını oluşturmaktadır. Bu durum, merkezi olmayan ancak senkronize çalışan YZ sistemlerinin toplumla daha uyumlu çalışabileceği yeni bir yönetim modeline de işaret etmektedir.

Ancak bu teknolojik ilerlemeler, etik ve toplumsal sorumluluk alanında da ciddi sorunları beraberinde getirmektedir. Makalede kapsamlı incelendiği gibi, Agentic AI sistemleri karar alma süreçlerinde şeffaf olmak zorundadır. “Neden bu kararı aldı?”, “Hangi verilerden etkilendi?” ve “Kararın sonucu kim tarafından üstlenilecek?” gibi soruların açık ve anlaşılır şekilde yanıtlanabilmesi, bu sistemlerin toplumsal kabulü açısından kritiktir. XAI bu noktada devreye girerek karar mekanizmalarının anlaşılabilirliğini ve güvenilirliğini artırmaktadır.

Bununla birlikte, büyük dil modellerinin eğitildiği veri setlerindeki önyargılar, Agentic AI sistemlerine de doğrudan yansımaktadır. Bu durum, toplumsal adalet ve eşitlik ilkeleriyle çelişebilecek kararların alınmasına neden olabilir. Özellikle sağlık, hukuk ve işe alım gibi hassas alanlarda önyargıların farkında olunması ve sürekli olarak bu sistemlerin denetlenmesi gerekmektedir. Ayrıca kullanıcı verileri açısından veri gizliliği ve siber güvenlik açısından da yeni riskler doğurmaktadır. Bu yüzden Agentic AI sistemlerinin etik ilkelere uygun, denetlenebilir ve şeffaf bir şekilde çalışmaları sağlanmalıdır.

Gelecekte, Agentic AI sistemleri yalnızca mevcut YZ teknolojilerini iyileştirmekle kalmayacak; aynı zamanda insan-merkezli karar alma sistemlerinin yerini alabilecek potansiyel bir aktör olarak karşımıza çıkacaktır. “Superintelligence” ve “Collective Intelligence” kavramlarıyla birlikte düşünüldüğünde, bu sistemlerin kendi hedeflerini belirleyebilme, kendi varlıklarını sürdürebilme ve çok sayıda ajanla ortak kararlar alabilme yetenekleri, klasik yapay zekâdan kopuşun en açık göstergeleridir.

Sonuç olarak bu çalışma, Agentic AI’nin yalnızca teknik bir inovasyon olmadığını; aynı zamanda toplumla, etik ve yönetimle entegre edilmesi gereken çok boyutlu bir dönüşüm süreci olduğunu ortaya koymuştur. Bu sistemlerin başarılı olabilmesi için, mühendislik yetkinliklerinin yanı sıra; etik reflekslerin, kullanıcı beklentilerinin, şeffaflık standartlarının ve hukuki çerçevelerin eş zamanlı geliştirilmesi gerekmektedir.

Agentic AI çağında, artık sadece “nasıl daha iyi bir sistem kurarız?” değil, aynı zamanda “bu sistemi kimin için, nasıl ve ne pahasına inşa ediyoruz?” soruları da her zamankinden daha fazla önem kazanmıştır.

Çizelge 1. Agentic AI, Genel AI ve CSI’nin özelliklerinin karşılaştırılması

Özellik	Agentic AI	General AI	Collective Superintelligence
Karar Verme	Otonom, hedef odaklı	İnsan benzeri genel düşünme	Grup etkileşimi ile yönlendirme
Hedef Belirleme	Kendi belirleyebilir.	Kullanıcı tarafından belirlenir.	Ortaklaşa şekillenir.
Öğrenme Yapısı	Pekiştirmeli, çok modlu	Derin öğrenme, çok görevli	YZ destekli etkileşim
Kullanıcı Rolü	İzleyici/denetleyici	Kullanıcı eğitir, girdi sağlar.	Aktif katılımcı
Açıklanabilirlik	Düşük-Orta	Orta	Yüksek
Toplumsal Uyum	Belirsiz, riskli	Etkileşim düzeyine bağlı	Katılımcı temelli, daha şeffaf
Öne Çıkan Risk	Alignment problemi, kontrol zayıflığı	Anlaşılmaz kararlar, zayıf güven	Yanılıcı yönlendirme, grup etkisi
Uygulama Alanı	Otonom sistemler, sağlık, finans	Genel YZ araştırmaları	Politikalar, karar destek sistemleri

Çizelge 2. Agentic AI, Genel AI ve CSI sistemleri

Kategori	Agentic AI	General AI	Collective Superintelligence
Performans	Yüksek görev uyumu	Geniş görev ilgisi	Kollektif doğruluk avantajı
Uyarlanabilirlik	Yüksek	Orta-Yüksek	Bağlama ve katılımcıya duyarlı
Etik Uyum	Belirsiz, riskli	Sınırlı kontrol	Etik yönetişime açık
Şeffaflık	Sınırlı	Kısmen mevcut	Yüksek
Kullanıcı Güveni	Düşebilir	Dalgalı	Yüksek etkileşimle artar.
Teknik Karmaşıklık	Orta	Yüksek	Orta-Yüksek

10. KAYNAKÇA

- Acharya, D. B., Kuppan, K. & Divya, B. (2025). Agentic AI: Autonomous intelligence for complex goals—A comprehensive survey. IEEE Access, 13. <https://doi.org/10.1109/ACCESS.2025.3532853>
- Hurlburt, G. (2017). Superintelligence: Myth or pressing reality? IEEE IT Professional, 19(1), 6–8. <https://doi.org/10.1109/MITP.2017.9>
- Ooge, J. & Verbert, K. (2021). Trust in prediction models: A mixed-methods pilot study on the impact of domain expertise. In 2021 IEEE Workshop on Trust and Expertise in Visual Analytics (Trex). <https://doi.org/10.1109/TREX53765.2021.00007>
- Rosenberg, L., Willcox, G. & Schumann, H. (2023). Towards collective superintelligence: A pilot study. In 2023 International Conference on Human-Centered Cognitive Systems (HCCS). <https://doi.org/10.1109/HCCS59561.2023.10452485>
- Balasubramani, R. & Biradar, V. G. (2024). Empowering Autonomous Decision-Making through Quantum Reinforcement Learning and Cognitive Neuromorphic Frameworks. In 2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC) (pp. 1–9). IEEE. <https://doi.org/10.1109/ICMNWC63764.2024.10872223>
- Vairaktarakis, G. & Vlachos, I. (2025). Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey. IEEE Access, 13, 18912–18951.
- Portugal, I. S., Alencar, P. & Cowan, D. (2024). An Agentic AI-based Multi-Agent Framework for Recommender Systems. 2024 IEEE International Conference on Big Data (Big Data), 5375–5381. <https://doi.org/10.1109/BIGDATA62323.2024.10825765>
- Stojanović, D., Pavković, B., Četić, N., Krnić, M. & Vidaković, L. (2024). Unit Test Generation Multi-Agent AI System for Enhancing Software Documentation and Code Coverage. 32nd Telecommunications Forum (TELFOR), 1–6. <https://doi.org/10.1109/TELFOR63250.2024.10819096>
- Gong, K., Gao, Y., Liang, X., Shen, X., Wang, M. & Lin, L. (2019). Graphonomy: Universal human parsing via graph transfer learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 7450–7459). IEEE. <https://doi.org/10.1109/CVPR.2019.00763>
- Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D. & Sikdar, B. (2023). A review of trustworthy and explainable artificial intelligence (XAI). IEEE Access, 11, 78994–79010. <https://doi.org/10.1109/ACCESS.2023.3294569>
- Sung, Y. H., Chien, S. Y. & Yu, F. (2024). Preliminary Validation of Explainable AI Interfaces across Heuristics and Information Transparency. Proceedings of the 4th IEEE International Conference on Human-Machine Systems (ICHMS). <https://doi.org/10.1109/ICHMS59971.2024.10555716>
- Kim, Y., Kim, M., Lee, J., Lee, H., Han, J. H. & Lee, W. H. (2024). Agentic LLM workflows for personalized user experience questionnaire generation. In 2024 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia) (pp. 1–6). IEEE. <https://doi.org/10.1109/ICCE-ASIA63397.2024.10773955>
- Fernando, R., Norton, I., Dogra, P., Sarnaik, R., Wazir, H., Ren, Z., Gunda, N. S., Mukhopadhyay, A. & Lutz, M. (2024). Quantifying bias in agentic large language models: A benchmarking approach. In 2024 5th Information Communication Technologies Conference (ICTC) (pp. 348–353). IEEE. <https://doi.org/10.1109/ICTC61510.2024.10601938>
- Rosenberg, L., Willcox, G. & Schumann, H. (2015). Towards collective superintelligence: A pilot study. In T. Lenaerts, M. Giacobini, H. Bersini, P. Bourguine, M. Dorigo & R. Doursat (Eds.), Proceedings of the European Conference on Artificial Life (ECAL 2015) (pp. 417–424). MIT Press.

- Kaindl, H. & Ferdigg, J. (2020). Towards an extended requirements problem formulation for superintelligence safety. In Proceedings of the IEEE Seventh International Workshop on Artificial Intelligence for Requirements Engineering (AIRE 2020) (pp. 1–6). IEEE.
- Ooge, J. & Verbert, K. (2021). Trust in prediction models: A mixed-methods pilot study on the impact of domain expertise. In 2021 IEEE Workshop on Trust and Expertise in Visual Analytics (TRES) (pp. 1–13). IEEE.
- Kamath, M. K., Mehta, S. S., Akshaya, H. L., Shilpashree, N., Jadhav, G., & Mitra, A. (2024). Neuromorphic-driven agentic AI for autonomous decision-making systems. 2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNBC), 1–8. IEEE. <https://doi.org/10.1109/ICARM57921.2024.10872131>
- Kshetri, N. (2025). Economics of agentic AI in the health-care industry. IT Professional. <https://doi.org/10.1109/MITP.2025.3529857>
- Singh, A., Ehtesham, A., Kumar, S. & Talaei Khoei, T. (2024). Enhancing AI Systems with Agentic Workflow Patterns in Large Language Models. In 2024 IEEE World AI IoT Congress (AllIoT) (pp. 1–6). IEEE. <https://doi.org/10.1109/AllIoT61789.2024.10578990>
- Ma, X. & Wang, J. (2024). WIP: Active learning through prompt engineering and agentic AI simulation – A pilot project in computer networks education. 2024 IEEE Frontiers in Education Conference (FIE). <https://doi.org/10.1109/FIE61694.2024.10892925>