



Contribution ID: 99

Tür: Oral Presentation

Yapay Zekâ Etiğinin Temellendirilmesi

20 Aralık 2024 Cuma 09:30 (30 dakika)

Yapay zekaların ahlaki bireyler olmaları nasıl sağlanabilir? Bu soru üzerine son yıllarda yürütülen çalışmalarda iki ana strateji ortaya çıkmaktadır. Birincisi yapay zekanın insan davranışını örnek almasını öngören “aşağıdan yukarı” yaklaşımı (Ng & Russell, 2000), diğeri ise yapay zekaya bazı temel deontik ilkelerin yüklenmesi gerektiğini öne süren “yukarıdan aşağı” yaklaşımı (Hooker, Kim & Davidson, 2021). Bu çalışmada her iki yaklaşımın da baş etmesi gereken temel bir sorunu ve bu soruna yönelik olası bir çözümü göstermeyi amaçlıyorum. Araştırmama yön veren soru şu şekilde ifade edilebilir: Gelecekte ahlaki bir birey olacak yapay zekanın ne tür “bilişsel” becerileri olmalıdır ki ahlaki ilkeleri tanıyabilsin ve kararlarında bu ilkelere yer verebilsin?

Yapay zekayı ne tür bir ahlaki yaklaşımla donatmamız gerektiği sorusuna yanıt ararken, yapay zekanın olası yetilerini ve kısıtlarını da hesaba katmak elzem olacaktır. Bunun için de, Kant'ın 1781 yılında yayınladığı meşhur eserinde kalktığı türden bir girişime, deyim yerindeyse Yapay Aklın Eleştirisi'ne yönelmek gerekecektir.

Kant felsefesinin anahtar kavramlarından birisi yargı gücüdür ve yargı gücünün kullanımı her daim bir refleksiyon içermelidir (Kant, 1998). Burada refleksiyon ile kast edilen husus bir yargıda bulunurken o yargının isabetli/doğru olduğundan nasıl emin olabileceğime ilişkin bir hassasiyettir. Bu araştırmada göstermeye çalışacağım husus şudur: gelecekte ahlaki bir birey olacak yapay zekanın bu türden bir hassasiyete sahip olması gerekecektir. Bu hassasiyetin hayata geçmesi de ancak “yukarıdan aşağı” bir yaklaşımla eşgüdümlü olarak mümkün olacaktır.

Anahtar Sözcükler: Yapay Zeka Etiği, Bilişsel Yetiler, Ahlaki Gerekçeleştirme, Deontoloji, Refleksiyon

E-mail / E-posta

ueldem@dogus.edu.tr

ORCID ID

0000-0003-2507-149X

Institution / Affiliation / Kurum

Doğuş Üniversitesi, Sosyoloji Bölümü

Country / Ülke

Türkiye

Presentation language / Sunum Dili

TR (Türkçe)

Disciplines / Disiplinler

Ethics / Etik

Başlıca yazarlar:: Umut Eldem (Doğuş Üniversitesi)

Sunu yapanlar: Umut Eldem (Doğuş Üniversitesi)

Session Classification: Session 1.6 (Day 3)

Track Classification: Congist'24: Theoretical Foundations