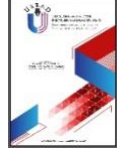




USBAD Uluslararası Sosyal Bilimler Akademi Dergisi -
International Journal of Social Sciences Academy, Yıl 3, Year
3, Sayı 5, Issue 5, Nisan 2021, April 2021
e issn: 2687-2641



GPT3: DALL-E VE JL2P EKSENİNDE VERİ GÖRSELLEŞTİRME VE HAREKETLENDİRME ÜZERİNE BİR İNCELEME

GPT3: DALL-E AND JL2P ON AXIS OF DATA VISUALIZATION AND MOTION
A REVIEW ON

Ezgi ŞEN

Güzel Sanatlar Enstitüsü, Grafik Anasanat Dalı, Anadolu Üniversitesi,
Eskişehir/Türkiye.

Fine Arts Institute, Graphic Major Art, Anadolu University,
Eskişehir/Turkey.

ezgisentr@gmail.com

ORCID ID: 000-0003- 2886-7098

Makale bilgisi | Article Information

DOI: 10.47994/usbadd.871726

Makale Türü / Article Type: Araştırma Makalesi / Research Article

Geliş Tarihi / Date Received: 31.01.2021

Kabul Tarihi / Date Accepted: 17.02.2021

Yayın Tarihi / Date Published: 20.04.2021

Yayın Sezonu / Pub Date Season: Nisan / April

Bu Makaleye Atıf İçin / To Cite This Article: Şen, E. (2021). GPT3: DALL-E ve JL2P Ekseninde Veri Görselleştirme ve Hareketlendirme Üzerine Bir İnceleme. *USBAD Uluslararası Sosyal Bilimler Akademi Dergisi* 3(5), 253-280.

İntihal: Bu makale intihal.net yazılımınca taranmıştır. İntihal tespit edilmemiştir.

Plagiarism: This article has been scanned by intihal.net. No plagiarism detected.



İletişim: Web: <https://dergipark.org.tr/tr/pub/usbadd>
mail: usbaddergi@gmail.com

Öz: Teknolojik evrimin güçlü sembollerinden biri olan yapay zekâ, derin öğrenme alanındaki güncel gelişmeler doğrultusunda hızla ilerleme kaydetmektedir. Bu ilerlemenin somut çıktılarında biri olan GPT3 sistemi (üretici ön-eğitilmiş dönüştürücü-generative pre-trained transformer3) derin öğrenme dil modeline bağlı olarak, veri görselleştirme alanında hızla gelişmektedir. OpenAI tarafından geliştirilen GPT3'ün bir ürünü olan Dall-e, doğal dilde ifade edilen metinleri görüntüye dönüştüren yapay bir sinir ağıdır. Dall'e 12 milyar civarında parametrenin anlamını tanıyarak kendisine verilen cümleleri daha önce hiç yapılmamış görsellere dönüştürebilmektedir. Dall-e sistemine benzer bir şekilde çalışan bir diğer sistem ise, JL2P (Joint Language to Pose) olarak adlandırılan doğal dile dayalı poz tahmini çalışmasıdır. JL2P uygulamasında tıpkı Dall-e de olduğu gibi yazı temelli bir aksiyon söz konusudur. Bu sistemde verilen sözel girdiye göre ilgili karakter, kendisinden istenen hareketi görsel olarak gerçekleştirebilmektedir. Bu durum animasyon ve oyun tasarımı alanında sadece sözel ifadeler üzerinden karakter görselleştirmelerinin ve hareketlerinin kolaylıkla yapılabileceğini göstermektedir. Bu makale, yapay zekânın veri görselleştirme ve hareketlendirme becerilerinin getirmiş olduğu güncel gelişmeleri, disiplinler arası bir bakış açısıyla okuyucuya aktarmayı amaçlamaktadır. Yapay zekâ ve veri görselleştirme alanındaki çalışmalar, tasarımcıların hikaye panosu (storyboard) oluştururken ya da beyin fırtınası (brainstorming) yaparken ilham almalarını sağlayabilecek bir sistemin kapılarını aralamaktadır. Bu alanda yapılan araştırmalar, tasarımcı ve yapay zekâ ilişkisinin yaratıcılık, üretkenlik, hız gibi kavramlar üzerinden değerlendirilmesi adına önemlidir.

Anahtar Kelimeler: Yapay zekâ, GPT3, GPT3 Dall-e, JL2P, Tasarım

Abstract: Artificial intelligence, one of the powerful symbols of technological evolution, progresses rapidly in line with current developments in the field of deep learning. One of the concrete outputs of this progress, the GPT3 system (producer pre-trained converter-generative pre-trained transformer3) is rapidly developing in the field of data visualization, depending on the deep learning language model. Dall-e, a product of GPT3 developed by OpenAI, is an artificial neural network that converts texts expressed in natural language into images. By recognizing the meaning of around 12 billion parameters to Dall, it can transform the sentences given to Dall into visuals that have never been done before. Another system that works similar to the Dall-e system is the natural language-based pose prediction study called JL2P (Joint Language to Pose). In the JL2P application, there is a text-based action just like Dall-e. According to the verbal input given in this system, the relevant character can perform the desired movement visually. This situation shows that character visualizations and movements can be easily done only through verbal expressions in the field of animation and game design. This article aims to convey to the reader the current developments brought about by the data visualization and activation skills of artificial intelligence with an

interdisciplinary perspective. Studies in the field of artificial intelligence and data visualization open the door to a system that can inspire designers while creating storyboards or brainstorming. Research in this field is important for evaluating the relationship between designer and artificial intelligence through concepts such as creativity, productivity, and speed.

Keywords: Artificial Intelligence, GPT3, GPT3 Dall-e, JL2P, Design

GİRİŞ

Teknolojik dönüşüm ve bilgi çağı, daha önce insanlar tarafından üstlenilen bazı rollerin yerini almaya başlayarak dijitalleşmenin varlığına işaret etmiştir. Son yıllarda ise, insanların iş ve sosyal yaşamının neredeyse her alanında yer almaya başlayan yapay zekâ, dijitalleşen dünyanın güçlü sembolleri arasında yerini almıştır. Schwab'a göre, dijitalleşme kavramı en genel tanımla otomasyona işaret etmektedir. Bu durum insan gücünden makine iş gücüne geçişi (otomasyonu) ortaya çıkarmıştır. Dijitalleşme ile artan bilgi-temelli ürünler ve teknolojik yenilikler, insan zekâsının değerini öne çıkartmıştır. İnsan zekâsının örnek alınarak oluşturulan yapay zekâ ise bu dönüşümün en somut örneklerinden biridir.

Yapay zekâ doğal dil kullanma, bellek, nesnelere tanıma gibi insan zekâsına özgü bazı özelliklerin biyolojik olmayan bir yapı içerisinde var olmasıdır. Yapay zekâ insanların yer aldığı birçok alanda (bankacılık, telekomünikasyon, kişisel asistanlar, eğitim, sağlık, savunma sanayi vb.) aktif olarak üretimler gerçekleştirmektedir. Bu alanların yanı sıra çizim, tasarım, görselleştirme, hareketlendirme gibi alanlarda da yapay zekânın birtakım gelişmeler göstermekte olduğu da bilinmektedir. Bu gelişmelerden en güncel ve dikkat çekici olanları, yapay zekâ ile veri görselleştirme ve hareketlendirme alanında yeni bir dönemin kapılarını aralayan GPT3 Dall-e ve JL2P'dir.

1. YAZI TEMELLİ VERİ GÖRSELLEŞTİRME: GPT3 DALL-E

Bu bölümde Dall-e'nin veri görselleştirirken oluşturduğu illüstrasyonlar ve farklı stillerde işlediği görüntülerin tasarım alanındaki etkisi, bir grafik tasarımcı bakış açısıyla değerlendirilip okuyucuya sunulmaktadır.

Sayısal ve sözel verilerin görselleştirilme tarihinin çok eskilere dayandığı bilinmektedir.

"Görselleştirmenin ilk örneklerinden olan geometrik diyagramlar, yıldızların konumu ve diğer ruhani değerler

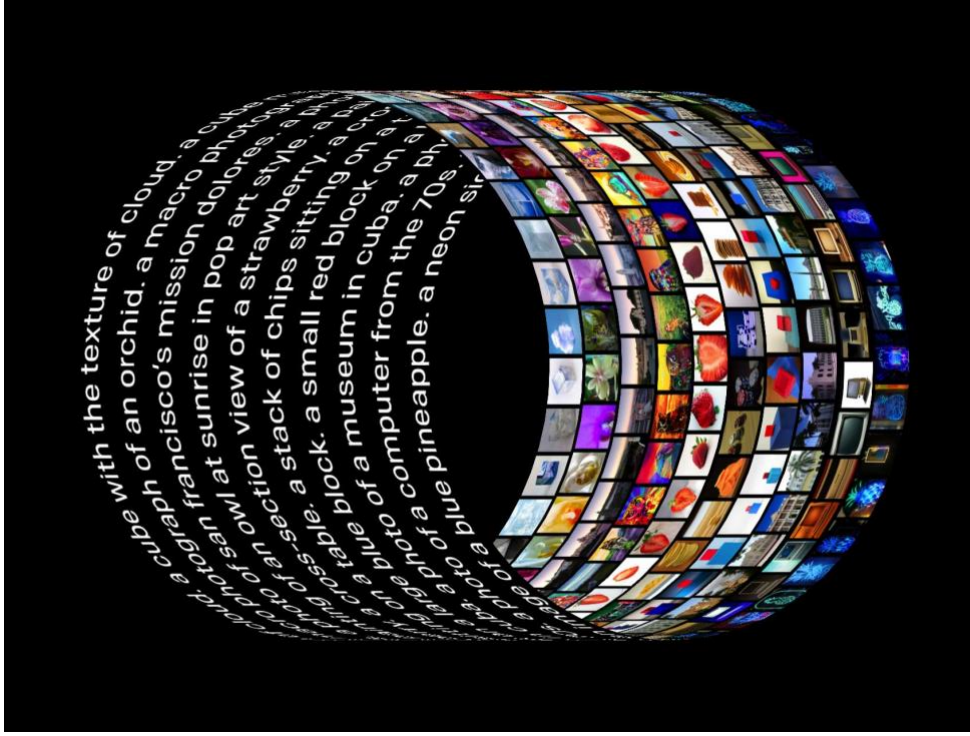
hakkında üretilen tablolar ile gerek yön bulma gerek keşfe çıkma amacıyla hazırlanan haritaların yapımında görülmüştür. İmgeleri doğrudan toparlamak ve matematik temelli tablolara kaydetmek konusunda da öncü girişimler görülmüştür. Bu girişimler, görselleştirme yönetiminin başlangıcını oluşturan ilk adımlardır” (Güler, 2008).

Veri görselleştirmesinin birçok alanla da etkileşimde olduğu açıktır.

“Veri görselleştirmesi, kendi gelişim süreci içerisinde farklı disiplinlerle etkileşim göstermekle beraber birçok görsel iletişim türüyle de etkileşim içerisinde olmuştur (Friendly, 2006). Veri görselleştirmesinin ilk adımının yazı olduğu söylenebilir. Yazı, İnsanların duygu ve düşüncelerini, bazı işaret ve sistemleri kullanarak meydana getirdikleri bir ifade biçimidir. Yazı, insanların yaşamının her alanında kullandığı bir iletişim biçimidir. Günümüzde bu iletişim biçimi, sayısal teknolojilerde bir bilgisayara veri görselleştirmesini sağlayabilecek konuma gelmeye başlamıştır. Makine öğrenmesine bağlı derin öğrenme alanındaki güncel gelişmeler, makinelerin çizim yapma, görselleştirme, stil kopyalama gibi yetilerinin hızla gelişmesine olanak sağlamaktadır. Makine öğrenmesi ‘Etkili ve hatasız tahmin yapabilen algoritmaların tasarlanmasından oluşmaktadır” (Mohri vd., 2012).

Bir diğer ifadeyle “Makine öğrenmesi alanı, deneyim ile otomatik olarak gelişen bilgisayar programlarının nasıl oluşturulabileceği ile ilgilenmektedir” (Mitchell, 1997). Örneklendirmek gerekirse, “Bir bilgisayara karşı Scrabble (veya başka bir oyun) oynadığınızı hayal edin. Başlangıçta her seferinde yenebilirsiniz, ancak birçok oyundan sonra sizi yenmeye başlar, ta ki sonunda asla kazanmayana kadar. Ya sen kötüleşiyorsun ya da bilgisayar Scrabble'da nasıl kazanılacağını öğreniyor. Sizi yenmeyi öğrendikten sonra, devam edip aynı stratejileri diğer oyunculara karşı kullanabilir, böylece her yeni oyuncuyla sıfırdan başlamaz; bu bir genelleme biçimidir (Zicari, 2018)”. Makine öğrenmesine bağlı derin öğrenme ise, "insan beyninin karmaşık problemler için gözlemeleme, analiz etme, öğrenme ve karar verme gibi yeteneklerini taklit etmeyi amaçlayan bir makine öğrenme sınıfıdır (Najafabadi vd., 2015)”. Bilgisayarlar derin öğrenmeyle kendi kendine öğrenebilir görüntü, ses ve metin gibi verileri anlamlandırıp soyutlayabilmektedir.

Derin öğrenme kullanılarak oluşturulan GPT3 Dall-e, yazı ile ifade ettiğimiz birçok cümleyi görselleştirebilen yapay bir sinir ağıdır.



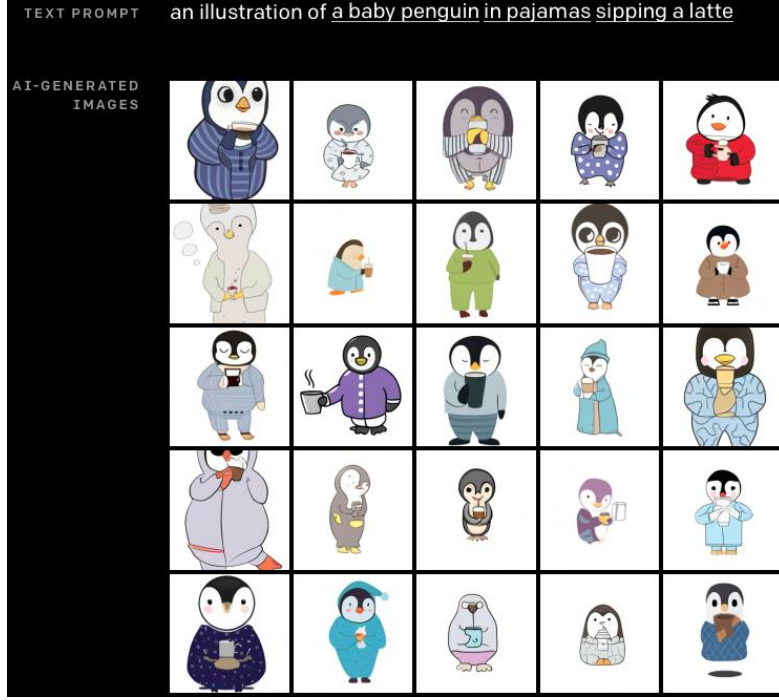
Görsel 1: Veri görselleştirme üzerine temsil görsel (OpenAI, 2021)

Dall-e olarak adlandırılması, Salvador Dali ve Pixar'ın Wall-e karakterinden ilham alarak oluşturulmasından kaynaklanmaktadır. GPT3 Dall-e metin ve görüntülerden oluşan veri seti kullanılarak, yazılardan görüntü oluşturmak üzere 12 milyar parametre ile yaratılan yapay bir sinir ağıdır (Openai, 2021). Dall-e, hayvanları ve nesnelere antropomorfize edebilir, metinleri mevcut görüntülere dönüştürebilir ve aynı sahnedeki bir görüntünün farklı stillerde işlenmesini sağlayabilmektedir. Dall-e çalışmasında kullanılan metin görüntü sentezinin temelleri 1983'lere kadar uzanmaktadır. "Metin-görüntü sentezi 1983 yılına kadar uzanıyor ve önemli bir evrime işaret ediyordu (Koh vd., 2020)" Üretken çekişmeli ağlardaki (GAN) gelişmeler uçtan uca eğitilebilir, foto-gerçekçi metin-görüntü üretimine olanak tanıyarak veri görselleştirme alanının gelişimine katkı sağlamıştır. Aşağıda GPT3 Dall-e tarafından ortaya konan illüstrasyonlar, farklı stil ve kompozisyondaki çizimler ve emojiler adım adım analiz edilerek okuyucuya sunulmuştur.

Yazı ve görselleştirme sisteminin çalışma biçimi aşağıda verilmiştir. Öncelikle, görselleştirilmek istenen konu ile ilgili bir cümlenin oluşturulması gerekmektedir.

Metin/Cümle: Pijamalı bebek penguen lattesini yudumlarırken.

Yapay zekâyla oluşturulan illüstrasyonlar:



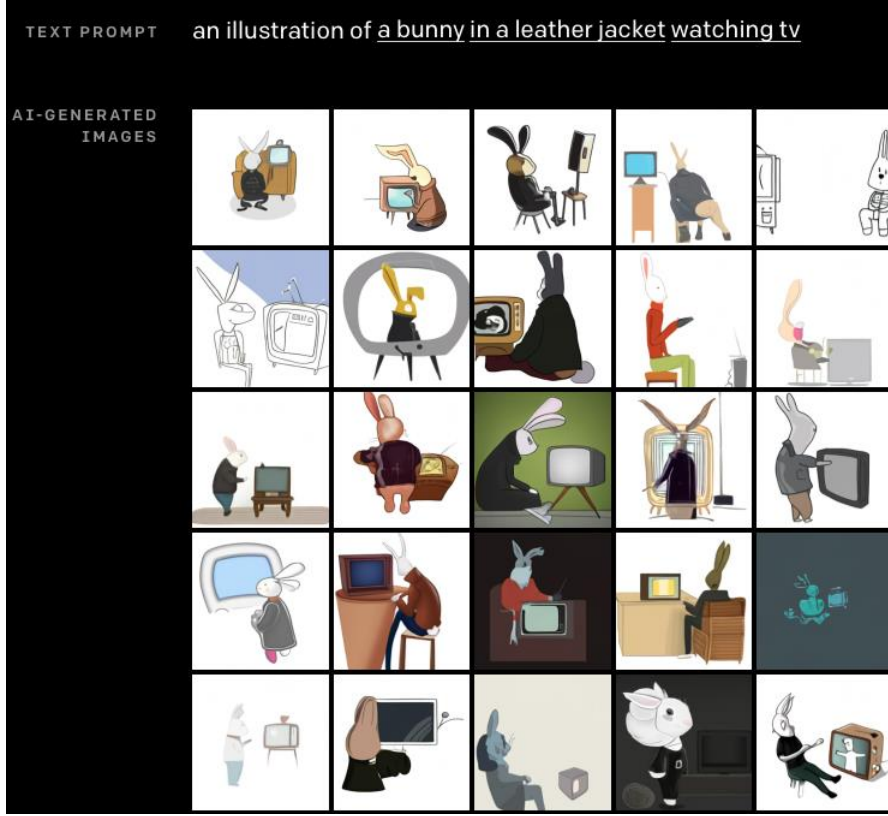
Görsel 2: Yapay Zekâ Tarafından Oluşturulan İllüstrasyonlar (Gpt3 dall-e, 2021)



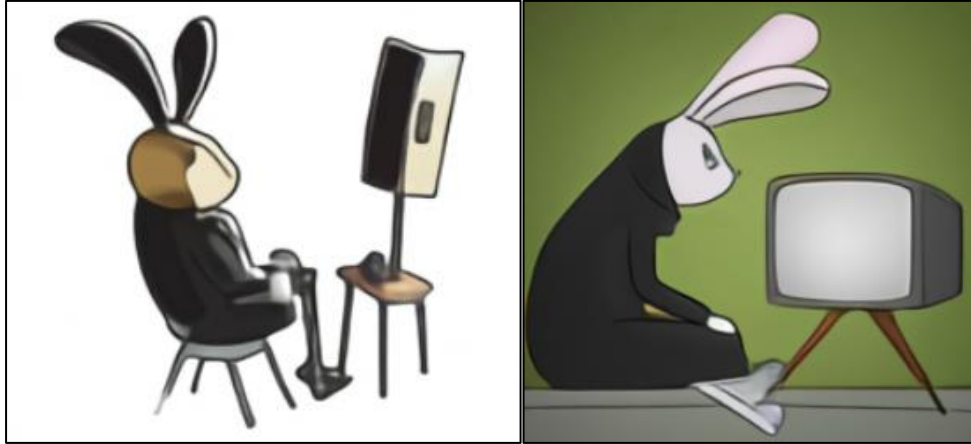
Görsel 3: Yapay Zekâ Tarafından Oluşturulan İllüstrasyonlar (Gpt3 dall-e, 2021)

Metin/Cümle: Deri ceketli bir tavşan televizyon izliyor.

Yapay zekâyla oluşturulan illüstrasyonlar:



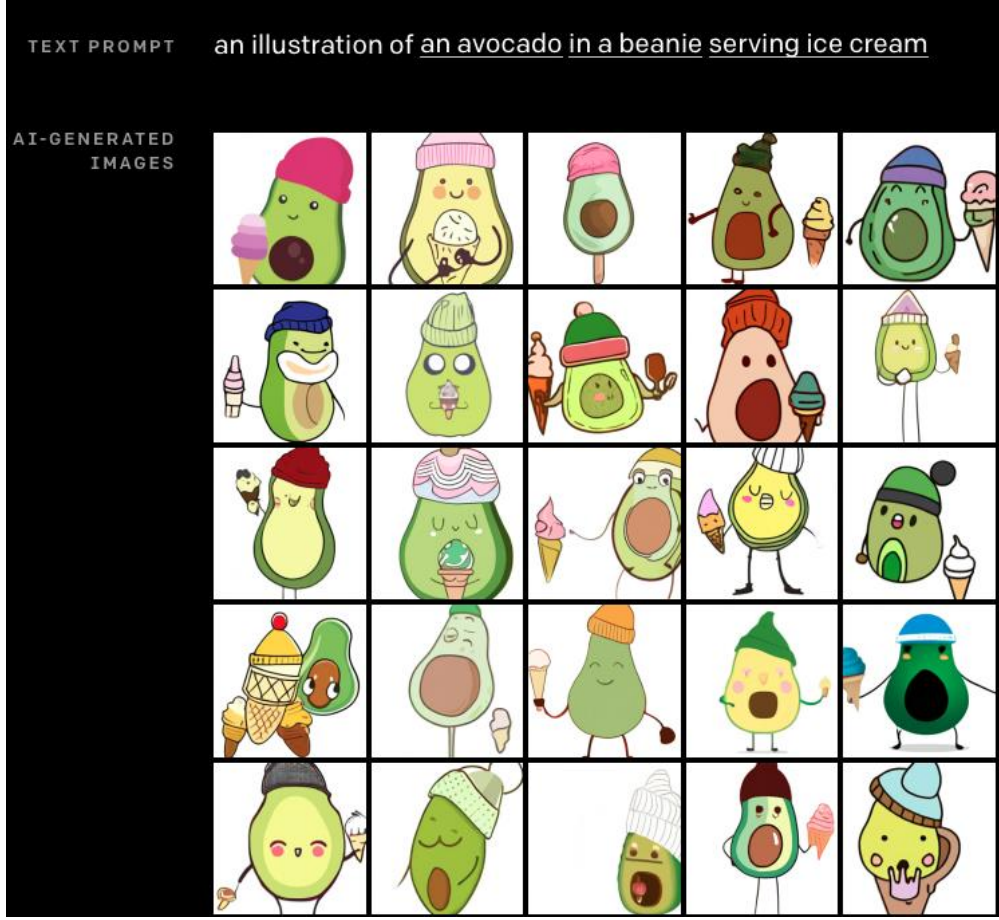
Görsel 4: Yapay Zekâ Tarafından Oluşturulan İllüstrasyonlar (Gpt3 dall-e, 2021)



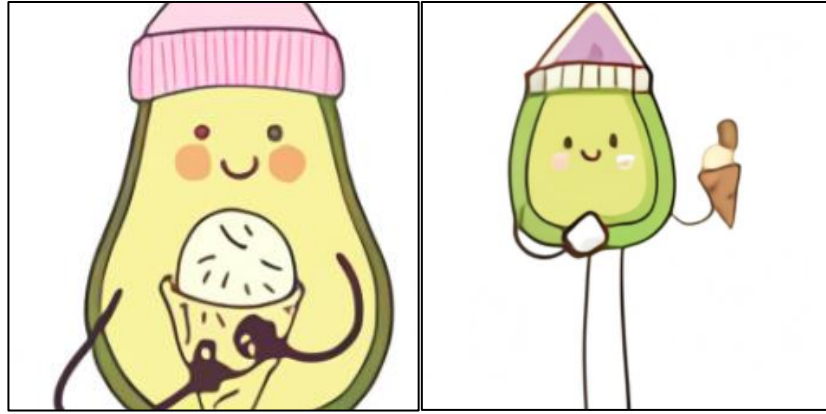
Görsel 5: Yapay Zekâ Tarafından Oluşturulan İllüstrasyonlar (Gpt3 dall-e, 2021)

Metin/Cümle: Bereli avokado dondurma alıyor.

Yapay zekâyla oluşturulan illüstrasyonlar:



Görsel 6: Yapay Zekâ Tarafından Oluşturulan İllüstrasyonlar (Gpt3 dall-e, 2021)

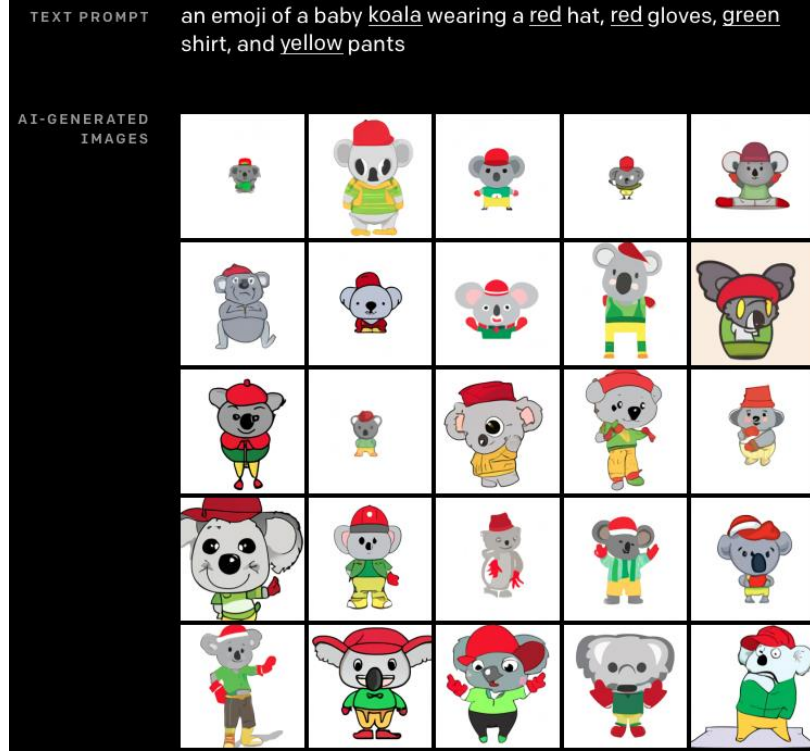


Görsel 7: Yapay Zekâ Tarafından Oluşturulan İllüstrasyonlar (Gpt3 dall-e, 2021)

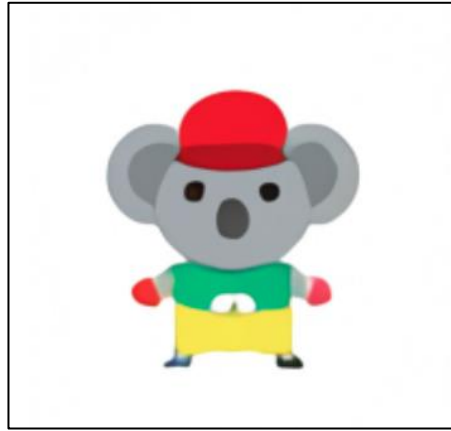
Sistem sadece cümleleri illüstrasyonlara çevirmekle kalmayıp ayrıca, renk kombinasyonlarını da doğru biçimde yerleştirebilmektedir.

Metin/Cümle: Bebek koala kırmızı şapka, kırmızı eldiven, yeşil t-shirt ve sarı pantolon giyiyor.

Yapay zekâyla oluşturulan görseller:



Görsel 8: Yapay Zekâ Tarafından Oluşturulan Renk Tanıma Çalışmaları (Gpt3 dall-e, 2021)

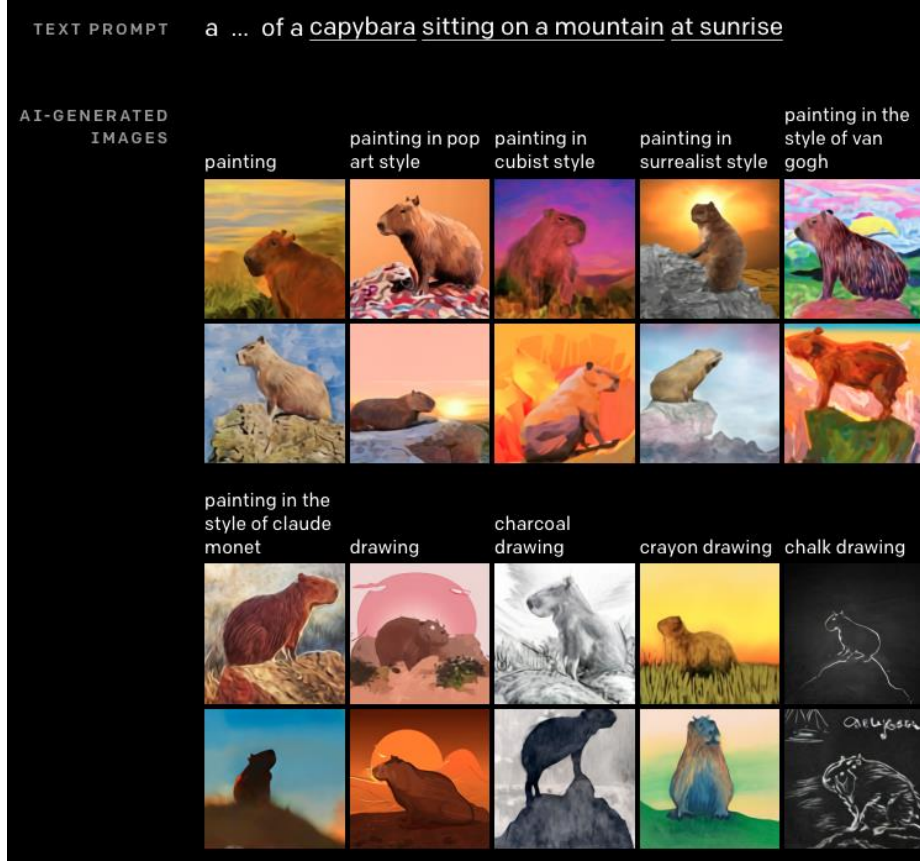


Görsel 9: Yapay Zekâ Tarafından Oluşturulan Renk Tanıma Çalışmaları (Gpt3 dall-e, 2021)

Dall-e verilen aynı sahneyi farklı stillerde işleyebildiği gibi ışık-gölge ve mevsime göre de tekrar görselleştirebilmektedir.

Metin/Cümle: Kapibara gün doğumunda dağda oturuyor.

Yapay zekâyla oluşturulan görseller:



Görsel 10: Yapay Zekâ Tarafından Stil, Işık/Gölge ve İklim'e Göre Oluşturulan Görseller (Gpt3 dall-e, 2021)

Çizimlerin farklı stil ve kompozisyondaki görünümleri

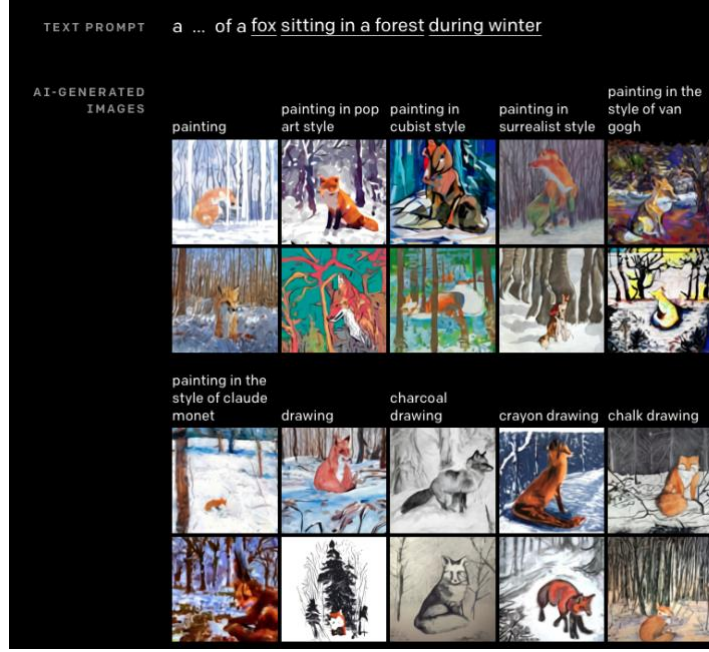


Görsel 11: Yapay Zekâ Tarafından Oluşturulan Stil Ve Kompozisyon Görselleri (Gpt3 dall-e, 2021)

Bahsedilen mevsime göre kompozisyonun tekrar tasarlanması.

Metin/Cümle: Bir tilki kış boyunca ormanda oturuyor.

Yapay zekâyla oluşturulan görseller:



Görsel 12: Yapay Zekâ Tarafından Oluşturulan Mevsime Göre Değişen Kompozisyon Görselleri (Gpt3 dall-e, 2021)

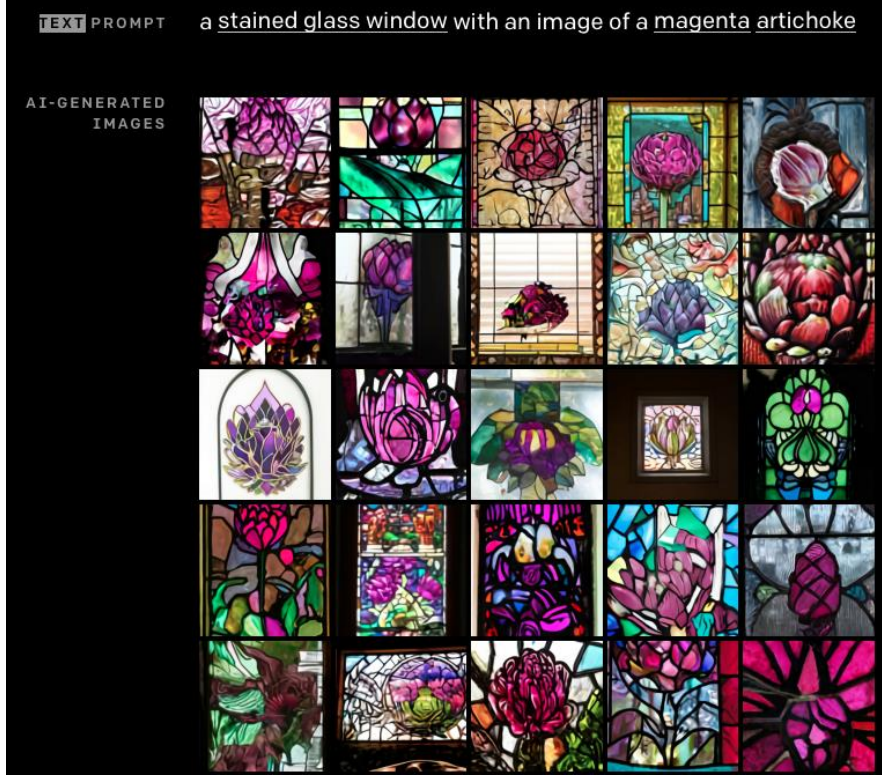


Görsel 13: Yapay Zekâ Tarafından Oluşturulan Mevsime Göre Değişen Kompozisyon Görselleri (Gpt3 dall-e, 2021)

Dall-e'nin bir diğer özelliği ise, verilen nesnenin kullanılacağı ortama göre (cam, şapka, t-shirt vb.) esnek bir şekilde illüstrasyonlarını uyarlamaya çalışmasıdır. Çizim yüzeyinin açısına bağlı olarak illüstrasyonlarını uyarlayabilmektedir. Aşağıda dall-e tarafından hazırlanan illüstratif enginar desenli vitray örnekleri bulunmaktadır. Burada da cümledeki özellik (renk, nesne, kullanılacak yüzey vb.) tanımlamalarına göre bir tasarlama süreci gerçekleştirilmiştir.

Metin/Cümle: Eflatun enginar resimli vitray pencere.

Yapay zekâyla oluşturulan görseller:



Görsel 14: Yapay Zekâ Tarafından Oluşturulan Vitray Görselleri (Gpt3 dall-e, 2021)

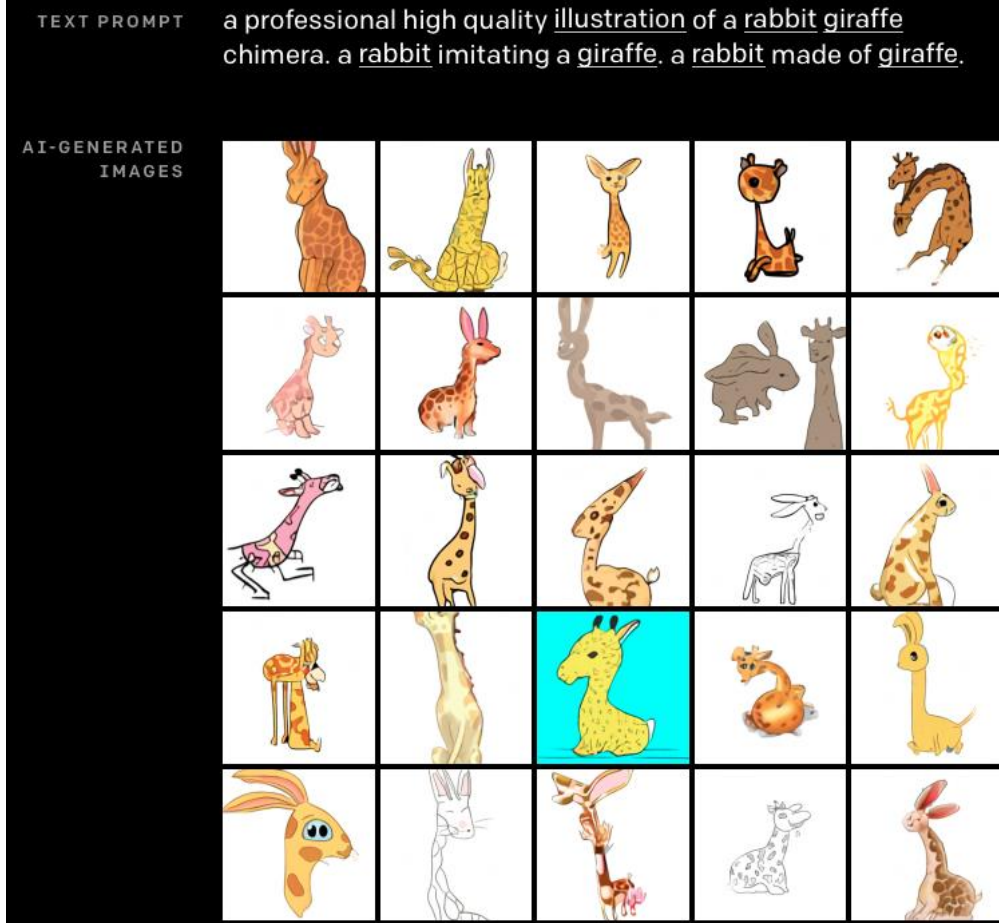


Görsel 15: Yapay Zekâ Tarafından Oluşturulan Vitray Görselleri (Gpt3 dall-e, 2021)

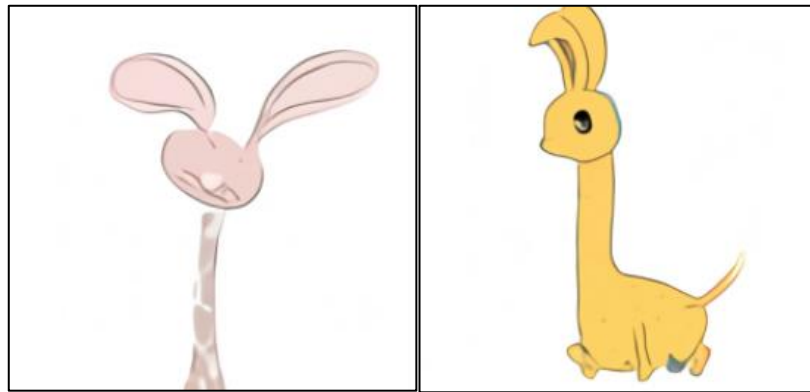
Dall-e farklı hayvanları birbiriyle ya da bir robotla birleştirip sıra dışı bir deneme gerçekleştirmektedir. Aşağıda bir tavşan ve zürafanın birleşiminde ortaya çıkan illüstrasyonlar bulunmaktadır.

Metin/Cümle: Yüksek çözünürlükte bir tavşan zürafa illüstrasyonu. Bir tavşan zürafayı taklit ediyor. Bir tavşan zürafa yapıldı.

Yapay zekâyla oluşturulan görseller:



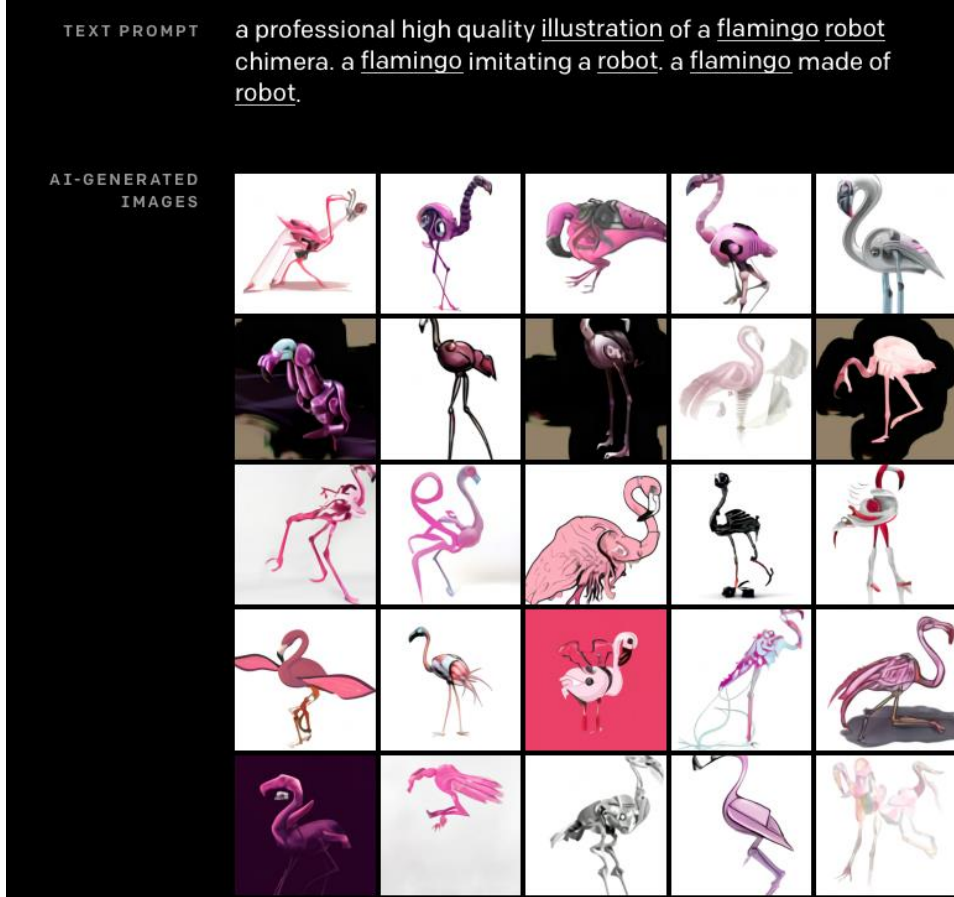
Görsel 16: Yapay Zekâ Tarafından Oluşturulan Hayvan Kombinasyonları (Gpt3 dall-e, 2021)



Görsel 16: Yapay Zekâ Tarafından Oluşturulan Hayvan Kombinasyonları (Gpt3 dall-e, 2021)

Metin/Cümle: Yüksek çözünürlükte bir robot flamingo illüstrasyonu. Bir flamingoyu taklit eden robot. Bir flamingo robot yapıldı.

Yapay zekâyla oluşturulan görseller: Dall-e tarafından oluşturulan flamingo görünümlü robot illüstrasyonları.



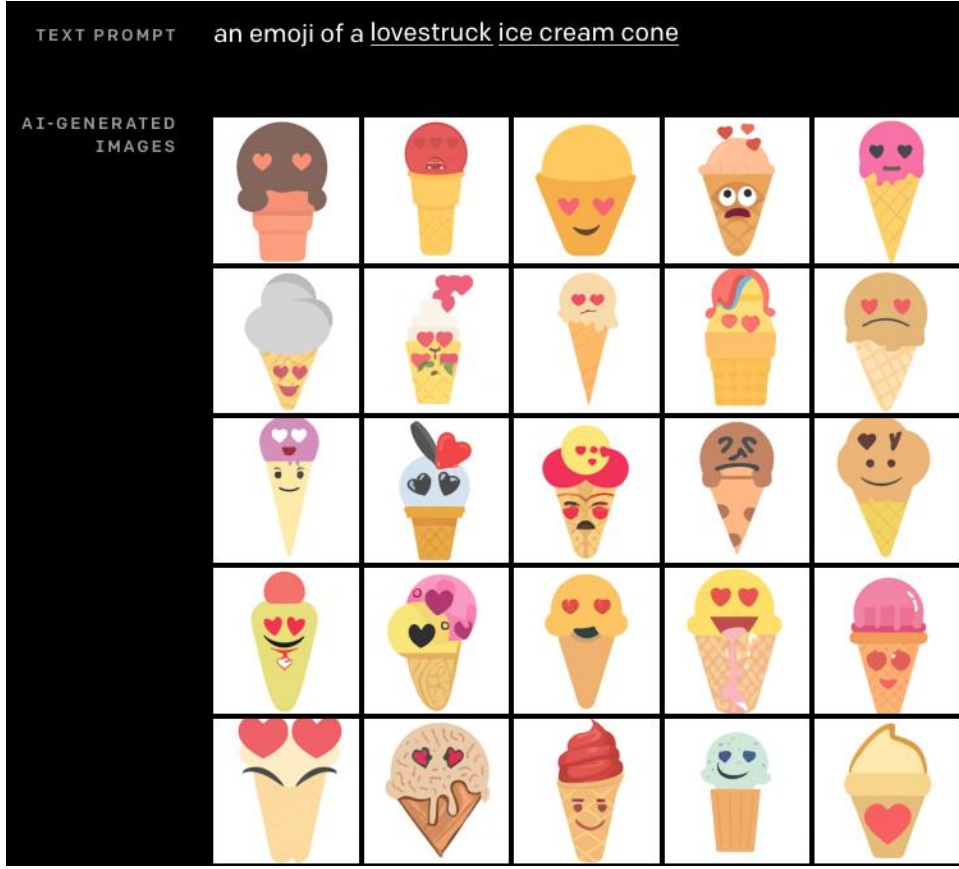
Görsel 17: Yapay Zekâ Tarafından Oluşturulan Robot Görselleri (Gpt3 dall-e, 2021)



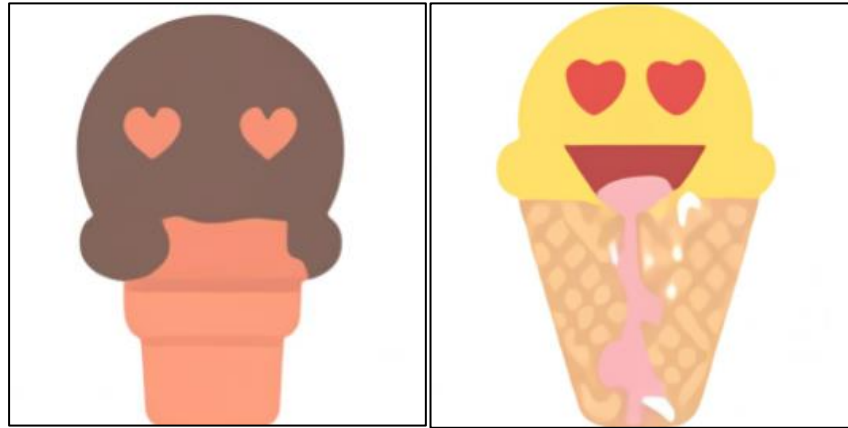
Görsel 18: Yapay Zekâ Tarafından Oluşturulan Robot Görselleri (Gpt3 dall-e, 2021)

Dall-e nin yaptığı bir diğer örnek ise emojileri canlı/cansız varlıklara (hayvan, yiyecek maddeleri vb.) aktarımı üzerinedir.

Metin/Cümle: Aşık olmuş bir dondurma külahı emoji.
Yapay zekâyla oluşturulan görseller:



Görsel 19: Yapay Zekâ Tarafından Oluşturulan Emoji Örnekleri (Gpt3 dall-e, 2021)



Görsel 20: Yapay Zekâ Tarafından Oluşturulan Emoji Örnekleri (Gpt3 dall-e, 2021)

Yukarıda GPT3 Dall-e uygulamaları üzerine birçok örnek verilmiştir. Bu örnekler yapay zekânın veri görselleştirmesine ilişkin geldiği noktayı tespit etmek adına önemlidir. Veri görselleştirilmesi tasarım disiplininin bir parçasıdır. İnsan zekâsının dışındaki bir varlığın bunu yapabiliyor olması, bu makalenin temel olarak incelediği bir

durumdur. Yukarıda verilen örneklerde dikkat edilen hususlar, yapay zekânın verilen cümleyi ve o cümledeki anahtar kelimeleri doğru anlayarak görsel çalışmalar üretebilmesidir. Yani doğru anatomi, renk, stil, kompozisyon gibi özelliklerin, gerçek tasarımcıların çalışmalarına yakın bir biçimde icra edilmesi oldukça önemlidir. İncelenen çalışmaların bazıları verilen cümlelere yakın ve hatta şaşırtıcı, ilham verici olarak seyrederken bazı çalışmaların ise verilen cümlenin dışında görseller ürettiği söylenebilir. Çalışmaların ne tam anlamıyla verilen cümlenin görselliğinden uzakta olduğu, ne de tamamen mükemmel olduğu açıktır. Örneğin verilen enginar detaylı eflatun vitray çalışması ilham verici ve şaşırtıcıyken, nesnelere emojilere dönüştürülmesi başarısız ve hatta sıradan olarak değerlendirilebilir.

2. AKILLI HAREKET TEKNOLOJİSİ: JL2P (Joint Language to Pose)

Bilgisayar bilimlerinde insanların kullandığı dil, *doğal dil işleme (NLP Natural Language Processing)* olarak önemli bir konumda yer almakta ve birçok proje doğal dilin referans alınmasıyla oluşturulmaktadır. Doğal dil kullanılarak oluşturulan tanımlamalar bir senaryonun görselleştirilip, hareketlendirilmesinin ilk adımıdır. Özellikle bir senaryo çerçevesinde hayat bulan canlandırma sineması ve belirli bir hareket planı içeren oyun tasarımı için hareket kavramı oldukça önemlidir. Geçmişte çoğunlukla el işi yöntemlerle hareketlendirilen karakterler, bilgisayar teknolojilerinin gelişmesiyle birlikte çeşitli programlar ve iş gücü (tasarımcı/operatörler) yardımıyla hareketlendirilmektedir. Son birkaç yılda ise, yapay zekâ alanındaki güncel gelişmeler doğrultusunda yazıyı (cümleleri) fiziksel hareketlere çevirebilen projeler geliştirilmeye başlanmıştır.

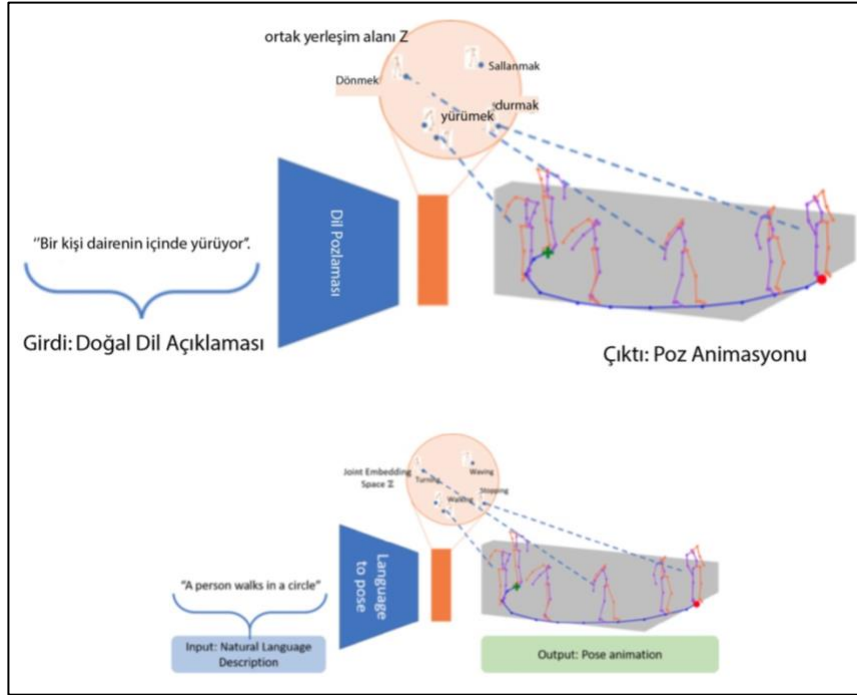
“Carnegie Mellon Üniversitesi'ndeki yapay zekâ araştırmacıları, Joint Language to Pose (JL2P) olarak adlandırılan doğal dili insan hareketlerine çevirerek fiziksel görevleri yerine getirebilen bir sistem geliştirmiştir” (bkz. Ahuja ve Morency, 2018).

Yazı temelli hareketlendirmelerde poz tahmini önemli bir yere sahiptir. Doğal dile dayalı poz tahmini, karakterin bir önceki pozundan bir sonraki poz dizisini anlamaya çalışıp oluşturmasıdır. Bu kısa süreli tahminler eller, bacaklar, baş ve gövdeye karşılık gelen eklem açılarını modellemeye odaklıdır. Uzun vadeli tahminlerde ise yürüme, koşma, zıplama ve emekleme gibi animasyonlar oluşturmak için karakterin

pozisyonları modellenmektedir. Resim veya konuşma koşullu poz tahminin de ise, insan referanslı görüntüler poz tahminin de daha sonra ne olacağını tahmin etmek için bir bağlam görevi görebilmektedir. Yüksek seviyeli bir kontrol parametresi olan bir görüntü, poz oluşturma için eylem bilgisine sahiptir, ancak hareket yörüngesinin hızı ve ivmesi üzerinde hassas bir kontrol sağlamamaktadır. Doğal dil cümleleri, eylemleri tanımlayan fiillerden, eylemin hızını/ivmesini tanımlayan zarflardan ve yönü veya hedefi tanımlayan sıfatlı isimlerden oluşmaktadır. Bu bilgiler, görüntü veya konuşmaya kıyasla poz üzerinde daha hassas bir kontrol sağlamaya yardımcı olabilmektedir. İnsan hareketleri genellikle kök ekleme bağlıdır, bu nedenle yörüngeyi öngörmek, doğal görünümlü animasyonlar elde etmek için önemlidir.

Girdi: (doğal dilde açıklaması) Bir insan dairenin içinde yürüyor.

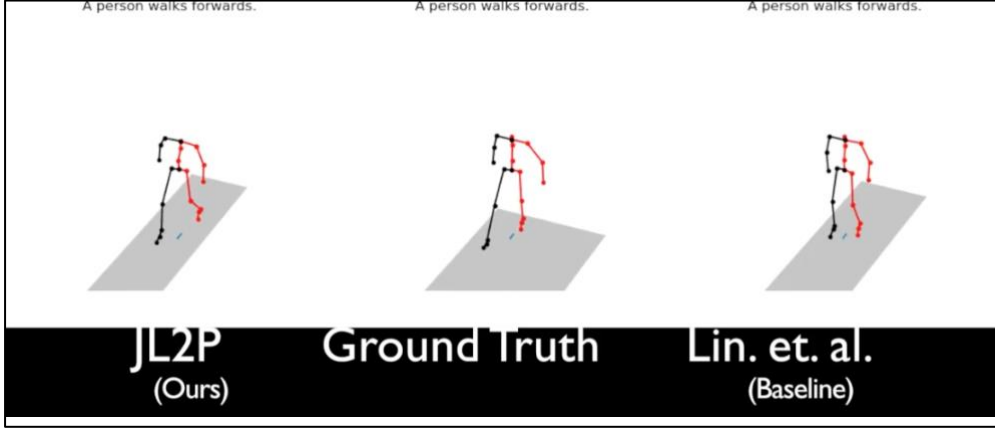
Çıktı: Poz animasyonu



Görsel 21: Doğal Dile Dayalı Poz Animasyonu Temsili Görseli (JL2P, 2021)

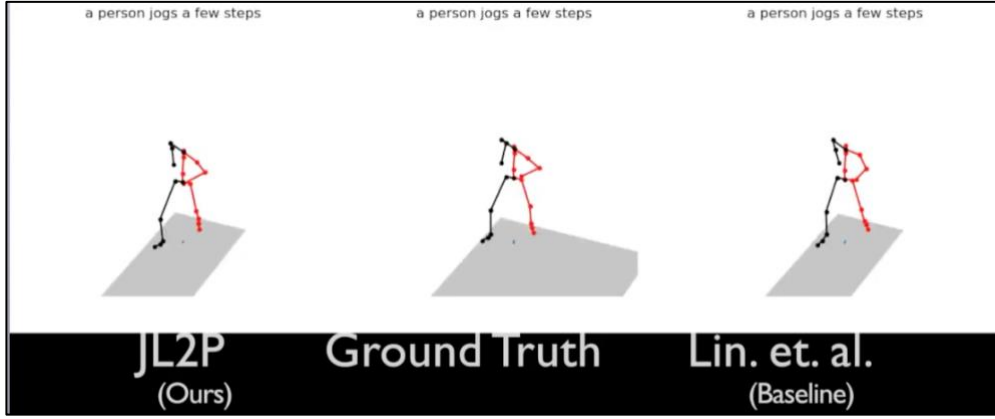
Aşağıda doğal dile dayalı poz tahminiyle yapılan animasyonlardan görüntüler yer almaktadır. Anlam bütünlüğünün sağlanması açısından <https://youtu.be/zfl1pcBIAVI> linkten hareketlerinin izlenmesi önemlidir.

Yön: Bir kişi ileri doğru yürüyor.



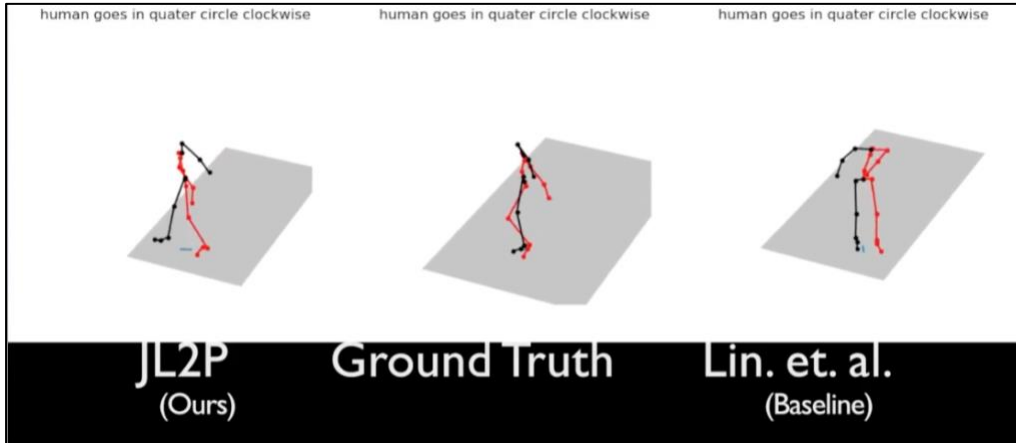
Görsel 22: Yön Kavramına Göre Yapılan Poz Tahmini Animasyonu (JL2P, 2021)

Hız: Bir kişi birkaç adım koşuyor.



Görsel 23: Hız Kavramına Göre Yapılan Poz Tahmini Animasyonu (JL2P, 2021)

Hareket: Saat yönünde çeyrek daire oluşturuyor.



Görsel 24: Hareket Kavramına Göre Yapılan Poz Tahmini Animasyonu (JL2P, 2021)

JL2P'nin becerileri arasında sağ ve sol gibi yön talimatlarını takip edebilme, yavaş veya hızlı olmak üzere hız kontrolü sağlayabilme gibi özellikler yer almaktadır. Biçimsel olarak çöp adamlarla denemeler

yapılsa da JL2P animasyonları, uzun vadede birçok tasarımın hareketlendirmesine katkı sağlayacaktır. JL2P gibi veri hareketlendirmesine yönelik çalışmaların oyun/animasyon gibi hareket temelli alanlara zaman, finans, iş gücü yönetimi gibi konularda katkı sağlayacağı açıktır.

TARTIŞMA VE SONUÇ

GPT3 Dall-e (veri görselleştirmesi) ve JL2P (veri hareketlendirmesi) projeleri geleceğe yönelik bir bakış açısıyla değerlendirildiğinde, yapay zekânın yazı ve görsel arasında doğru bir ilişki kurmaya çalışarak nispeten anlamlı görseller çıkarabildiğini göstermektedir. Hareketlendirme alanında ise, yapay zekânın belirli seviyede başarılı olduğu ve alandaki çalışmaların gelişerek devam ettiği söylenebilir. Veri görselleştirme ve hareket alanındaki bu gelişmeler tasarımcıların:

- Fikir bulma aşamasında hızlı eskizler yapabilmelerini sağlayacaktır.
- Planladıkları hareketi karakter üzerinde eş zamanlı görebilmelerini sağlayacaktır.
- Mevcut fikirlerini hikaye panosuna dökerek hızlı bir tasarım planı yapmalarını sağlayacaktır.
- İş gücünü ve zaman kullanımını olumlu yönde etkileyecektir.
- Yapay zekânın ortaya koyduğu illüstrasyonlardan ilham alarak yaratıcı ve özgün çalışmalar yapmalarına imkan sağlayacaktır.
- Tek bir cümle ile sıra dışı çizim, renk, kompozisyon denemeleri yapmalarını sağlayacaktır.
- Günümüzde kullanılan tasarım programlarına göre projelerde daha hızlı sonuçlar elde edilmesini sağlayacaktır.

Yapay zekânın tasarım yapabilme becerisi, bu zamana kadar yapay zekânın bir hesaplama aracından yaratıcı bir araca dönüşümünü göstermektedir. Makale kapsamında bahsi geçen iki teknolojinin bileşimi, yakın bir zamanda yapay zekânın kendi karakterlerini hem tasarlaması hem de hareketlendirmesi olasılığını artırmaktadır. Tasarımda henüz yeni bir aktör olan yapay zekânın günümüz tasarımcılarına alternatif bir araç olması, yarının tasarım dünyasında iş gücü, maliyet ve üretim bakımından ciddi bir rekabet potansiyelini taşımaktadır.

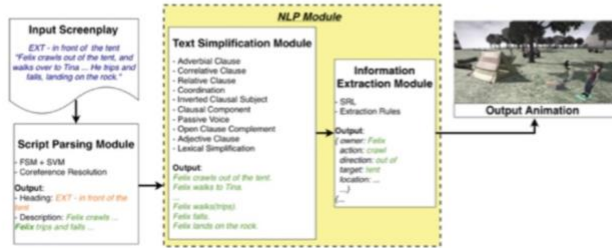
Tüm bu çalışmalara ek olarak; GPT3 Dall-e'nin veri görselleştirme alanında yeni olmadığı söylenebilir. Dall-e'nin (2021) yanı sıra

Microsoft ve JD AI Labs araştırmacılarının (2019) tanıttığı ObjGAN (nesneye dayalı üretken çekişmeli ağlar) tanımlayıcı cümlelerle görüntüler üretebilmektedir.



Görsel 25: Microsoft ObjGAN Veri Görselleştirme Sistemi, 2019

Bir diğer çalışma ise, Disney'in ürettiği (2019) yapay zekâ algoritmasıdır. "Disney Research ve Rutgers Üniversitesi bilim insanları, hikaye panosu (storyboard) ve video oluşturmak için yapay zekâ kullanarak bir model geliştirmiştir" (bkz. Zhang vd., 2019). Sistem, hikaye panosu (storyboard) üretmek için senaryodaki metinlerden faydalanmaktadır.



Görsel 26: Disney Veri Görselleştirme Sistemi, 2019

Dall-e veri görselleştirme alanındaki en yeni çalışma olarak gösteriyor ki, yapay zekâ hızla kendini geliştirerek daha anlamlı sonuçlar ortaya koymaya başlamaktadır. Yazı kullanılarak bir verinin görselleştirilmesi/hareketlendirilebilmesi grafik, animasyon, oyun

tasarımı alanında üretimler gerçekleştiren herkesin yakın gelecekte kullanabileceği bir tasarım aracı haline gelecektir.

KAYNAKÇA

- Ahuja, C. & Morency, L. P. (2018). Language2Pose: Natural Language Grounded Pose Forecasting. Cornell University, Erişim Adresi: <https://arxiv.org/abs/1907.01108>.
- Friendly, M. & Brief, A. (2006). *History of Data Visualization: Handbook of Computational Statistics*. Toronto: Canada.
- Güler, T. (2008). Bilgilendirme Tasarımı. *Grafik Tasarım Dergisi* 21, 28.
- Koh, Y. J., Baldrige, J., Lee, H. & Yang Y. (2020). *Text-to-Image Generation Grounded by Fine-Grained User Attention*. Google Research. Erişim Adresi: <https://arxiv.org/pdf/1711.10485.pdf>
- Schwab, K. (2017). *Dördüncü Sanayi Devrimi*. İstanbul: Optimist Yayınları.
- Mohri, M., Rostamizadeh, A. & Talwalkar, A. (2012). *Foundations of Machine Learning*. Cambridge: The MIT Press.
- Mitchell, T. M. (1997). *Machine Learning*. 1st Edition. New York: McGraw-Hill.
- Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R. & Muharemagic, E. (2015). Deep Learning Applications and Challenges in Big Data Analytics. *Journal of Big Data* 2(1), 1–21.
- Zhang, Y., Tshipidi, E., Schriber, S., Kapadia, M., Gross, M. & Modi, A. (2019). Generating Animation from Screenplays. Disney Research&Rutgers University. Erişim Adresi: <https://arxiv.org/pdf/1904.05440.pdf>.
- Zicari, V. R. (2018). *Explorations in Artificial Intelligence and Machine Learning*. Florida: CRC Press.
- OpenAI. (2021). Erişim Adresi: <https://openai.com/blog/organizational-update/> Erişim Tarihi: 14.01.2021.
- Openai. (2021). Erişim Adresi: <https://youtu.be/zflipcBlAVI> Erişim Tarihi: 15.01.2021.

GÖRSEL LİSTESİ

Görsel 1: Veri görselleştirme üzerine temsil görsel. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 2: Yapay zekâ tarafından oluşturulan illüstrasyonlar. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 3: Yapay zekâ tarafından oluşturulan illüstrasyonlar. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 4: Yapay zekâ tarafından oluşturulan illüstrasyonlar. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 5: Yapay zekâ tarafından oluşturulan illüstrasyonlar. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 6: Yapay zekâ tarafından oluşturulan illüstrasyonlar. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 7: Yapay zekâ tarafından oluşturulan illüstrasyonlar. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 8: Yapay zekâ tarafından oluşturulan renk tanıma çalışmaları. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 9: Yapay zekâ tarafından stil, ışık/gölge ve iklime göre oluşturulan görseller. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 10: Yapay zekâ tarafından stil, ışık/gölge ve iklime göre oluşturulan görseller. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 11: Yapay zekâ tarafından oluşturulan stil ve kompozisyon görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 12: Yapay zekâ tarafından oluşturulan mevsime göre değişen kompozisyon görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 13: Yapay zekâ tarafından oluşturulan mevsime göre değişen kompozisyon görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 14: Yapay zekâ tarafından oluşturulan vitray görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 15: Yapay zekâ tarafından oluşturulan vitray görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 16: Yapay zekâ tarafından oluşturulan hayvan kombinasyonları. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 17: Yapay zekâ tarafından oluşturulan robot görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 18: Yapay zekâ tarafından oluşturulan robot görselleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 19: Yapay zekâ tarafından oluşturulan emoji örnekleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 20: Yapay zekâ tarafından oluşturulan emoji örnekleri. (2021). Erişim adresi: <https://openai.com/blog/dall-e/>

Görsel 21: Doğal dile dayalı poz animasyonu temsili görseli. (2019). Erişim adresi: <https://arxiv.org/pdf/1907.01108.pdf>

Görsel 22: Yön kavramına göre yapılan poz tahmini animasyonu. (2019). Erişim adresi: <https://arxiv.org/pdf/1907.01108.pdf>

Görsel 23: Hız kavramına göre yapılan poz tahmini animasyonu. (2019). Erişim adresi: <https://arxiv.org/pdf/1907.01108.pdf>

Görsel 24: Hareket kavramına göre yapılan poz tahmini animasyonu. (2019). Erişim adresi: <https://arxiv.org/pdf/1907.01108.pdf>

Görsel 25: Microsoft ObjGAN veri görselleştirme sistemi. (2019). Erişim adresi: <https://medium.com/syncedreview/microsoft-obj-gan-turns-words-into-complex-scenes-5c6024f0f91d>

Görsel 26: Disney veri görselleştirme sistemi. (2019). Erişim adresi: <https://studios.disneyresearch.com/wp-content/uploads/2019/08/Generating-Animations-from-Screenplays.pdf>

GENİŞLETİLMİŞ ÖZET: Teknolojik dönüşüm ve bilgi çağı, daha önce insanlar tarafından üstlenilen birçok rolün değişmesini sağlamış ve dijitalleşmeye zemin hazırlamıştır. Son yıllarda ise, insanların iş ve sosyal yaşamının neredeyse her alanında yerini almaya başlayan yapay zekâ, dijitalleşen dünyanın güçlü sembolleri arasında yerini almıştır. Yapay zekânın, derin öğrenme alanındaki güncel gelişmeler doğrultusunda hızla ilerleme kaydettiği söylenebilir. Derin öğrenme bilgisayarların var olan verilerle (yapay sinir ağları ve algoritmalar kullanarak) yeni veriler elde edebilmesidir. Aynı zamanda makine öğrenmesinin de bir alt kümesidir. Derin öğrenme alanındaki ilerlemenin somut çıktılarında biri olan GPT3 sistemi (üretici ön-eğitilmiş dönüştürücü-generative pre-trained transformer3) derin öğrenme dil modeline bağlı olarak, veri görselleştirme alanında hızla gelişim kaydetmektedir. Gpt3 insanların duygu ve düşüncelerini ifade ettiği

metin/cümlelerin görselleştirilmesini sağlayan derin öğrenme tabanlı bir dil modelidir. GPT3'ün bir parçası olan Dall-e, doğal dilde ifade edilen metinlerden görüntüler oluşturabilen yapay bir sinir ağıdır. OpenAI tarafından geliştirilen GPT3 Dall-e, yazı ile ifade ettiğimiz birçok cümleyi görselleştirmeye kalmayıp, bunu farklı stil ve kompozisyonlarda yapabilmektedir. Dall-e olarak adlandırılmasının sebebi, Salvador Dalí ve Pixar'ın Wall-e karakterinden ilham alarak oluşturulmasından kaynaklanmaktadır. GPT3 Dall-e metin ve görüntülerden oluşan veri setleri kullanılarak, yazılardan görüntü oluşturmak üzere 12 milyar parametre ile yaratılan yapay bir sinir ağıdır. Dall-e, hayvanları ve nesnelere antropomorfize edebilir, metinleri mevcut görüntülere dönüştürebilir ve aynı sahnedeki bir görüntünün farklı stillerde işlenmesini sağlayabilmektedir. Yapay zekâ ve veri görselleştirme alanındaki çalışmalar, tasarımcıların hikâye panosu (storyboard) oluştururken ya da beyin fırtınası (brainstorming) yaparken ilham almalarını sağlayabilecek bir sistemin kapıları aralamaktadır. Veri görselleştirme alanındaki araştırmalar tasarımcı ve yapay zekâ ilişkisini yaratıcılık, üretkenlik, hız gibi kavramlar üzerinden değerlendirilmesine olanak sağlaması adına önemlidir. Dall-e sistemine benzer bir şekilde çalışan bir diğer sistem ise, JL2P (Joint Language to Pose) olarak adlandırılan doğal dile dayalı poz tahmini çalışmasıdır. JL2P uygulamasında Dall-e de olduğu gibi yazı temelli bir aksiyon söz konusudur. Carnegie Mellon Üniversitesi'ndeki yapay zekâ araştırmacıları, Joint Language to Pose (JL2P) olarak adlandırılan doğal dili insan hareketlerine çevirerek fiziksel görevleri yerine getirebilen bir zekâ sistemi geliştirmiştir. Yazı temelli hareketlendirmelerde poz tahmini önemli bir yere sahiptir. Doğal dile dayalı poz tahmini, karakterin bir önceki pozundan bir sonraki poz dizisini anlamaya çalışıp oluşturmasıdır. Bu kısa süreli tahminler eller, bacaklar, baş ve gövdeye karşılık gelen eklem açılarını modellemeye odaklıdır. Uzun vadeli tahminlerde ise yürüme, koşma, zıplama ve emekleme gibi animasyonlar oluşturmak için karakterin pozisyonları modellenmektedir. Resim veya konuşma koşullu poz tahmininin de ise, insan referanslı görüntüler poz tahmininin de daha sonra ne olacağını tahmin etmek için bir bağlam görevi görebilmektedir. Yüksek seviyeli bir kontrol parametresi olan bir görüntü, poz oluşturma için eylem bilgisine sahiptir ancak hareket yörüngesinin hızı ve ivmesi üzerinde hassas bir kontrol sağlamamaktadır. Doğal dil cümleleri, eylemleri tanımlayan fiillerden, eylemin hızını/ivmesini tanımlayan zarflardan ve yönü/hedefi tanımlayan sıfatlı isimlerden oluşmaktadır. Bu bilgiler, görüntü veya konuşmaya kıyasla poz üzerinde daha hassas bir kontrol sağlamaya yardımcı olabilmektedir. İnsan hareketleri genellikle kök eklemeye bağlıdır, bu nedenle yörüngeyi öngörmek, doğal görünümlü animasyonlar elde etmek için önemlidir. Bu sistem ile ilgili kişi, bir karaktere ya da nesneye istediği hareketi yazılı olarak verdiğinde karakter hareketi yapabilmekte, yazı harekete çevrilebilmektedir. Bu durum animasyon ve oyun tasarımı alanında sadece cümlelerle karakter

hareketlerinin kolaylıkla yapılabilineceğini göstermektedir. Yapay zekâ alanındaki güncel gelişmelerle, yalnızca yazı kullanılarak bir verinin görselleştirilip hareketlendirilmesi mümkün hale gelmeye başlamıştır. GPT3 Dall-e (veri görselleştirmesi) ve JL2P (veri hareketlendirmesi) katastrofik senaryolardan uzak geleceğe yönelik bir bakış açısıyla değerlendirildiğinde, yapay zekânın yazı ve görsel arasında doğru bir ilişki kurmaya çalışarak nispeten anlamlı görseller çıkarabildiğini göstermektedir. Hareketlendirme alanında ise, yapay zekânın belirli seviyede başarılı olduğu ve alandaki çalışmaların gelişerek devam ettiği söylenebilir. Veri görselleştirmesi, tasarımcıların fikir bulma aşamasında hızlı eskizler yapabilmelerine olanak sağlayacaktır. Aynı zamanda, var olan fikirlerini hikaye panosuna dökerek hızlı bir tasarım planı yapmalarını sağlayacaktır. Veri hareketlendirmesi, planlanan hareketin karakter üzerinde eş zamanlı görülebilmesini sağlayacaktır. İş gücünü ve zaman kullanımını olumlu yönde etkileyecektir. Yapay zekânın ortaya koyduğu illüstrasyonlar tasarımcılara ilham vererek yaratıcı ve özgün çalışmalar yapmalarına imkan sağlayacaktır. Tek bir cümle ile sıra dışı çizim, renk, kompozisyon denemeleri yapılabilecektir. Günümüzde kullanılan tasarım programlarından biri olarak projelerde daha hızlı sonuçlar elde edilmesini sağlayacaktır. Yapay zekâ hızla kendini geliştirerek daha anlamlı sonuçlar ortaya koymaya başlamaktadır. Yazı kullanılarak bir verinin görselleştirilmesi/hareketlendirilebilmesi grafik, animasyon, oyun tasarımı alanında üretimler gerçekleştiren herkesin yakın gelecekte kullanabileceği bir tasarım aracı haline gelecektir. Bu makale, yapay zekânın veri görselleştirme ve hareket sistemine getirmiş olduğu güncel gelişmeleri, disiplinlerarası bir bakış açısıyla okuyucuya aktarmayı amaçlamaktadır.

EXTENDED ABSTRACT: Technological transformation and the information age have changed many roles previously assumed by people and paved the way for digitalization. In recent years, artificial intelligence, which has started to take its place in almost every aspect of people's business and social life, has taken its place among the powerful symbols of the digitalizing world. It can be said that artificial intelligence has made rapid progress in line with current developments in deep learning. Deep learning is the ability of computers to obtain new data with existing data (using artificial neural networks and algorithms). It is also a subset of machine learning. The GPT3 system (producer pre-trained transformer-generative pre-trained transformer3), which is one of the concrete outputs of the progress in the field of deep learning, rapidly develops in the field of data visualization depending on the deep learning language model. Gpt3 is a deep learning-based language model that enables visualization of texts/sentences in which people express their feelings and thoughts. Dall-e, part of GPT3, is an artificial neural network that can generate images from texts expressed in natural language. Developed by OpenAI, GPT3 Dall-e not only visualizes many sentences we express in the text but can do this in different styles and

compositions. The reason it is called Dall-e is that it was created inspired by the character Wall-e by Salvador Dalí and Pixar. GPT3 Dall-e is an artificial neural network created with 12 billion parameters to create images from text using data sets consisting of text and images. Dall-e can anthropomorphize animals and objects, transform texts into existing images, and render an image in the same scene in different styles. Studies in the field of artificial intelligence and data visualization open the door to a system that can inspire designers while creating storyboards or brainstorming. Research in the field of data visualization is important in order to allow the relationship between designer and artificial intelligence to be evaluated through concepts such as creativity, productivity, and speed. Another system that works similarly to the Dall-e system is the natural language-based pose estimation study called JL2P (Joint Language to Pose). In the JL2P application, there is a text-based action as in Dall-e. Artificial intelligence researchers at Carnegie Mellon University have developed an intelligent system that can perform physical tasks by translating natural language into human movements, called the Joint Language to Pose (JL2P). Pose estimation has an important place in text-based movements. Pose prediction based on natural language is when the character tries to understand and construct the next pose from the previous pose. These short-term estimates focus on modeling the joint angles corresponding to the hands, legs, head, and torso. In long-term predictions, the positions of the character are modeled to create animations such as walking, running, jumping, and crawling. As for picture or speech conditional exposure estimation, human referenced images can also serve as a context for predicting what will happen next. An image with a high level of control parameter has action information for posing but does not provide precise control over the velocity and acceleration of the motion trajectory. Natural language sentences consist of verbs that describe actions, adverbs that describe the speed/acceleration of action, and adjective nouns that describe the direction/goal. This information can help provide more precise control of the pose compared to image or speech. Human movements often depend on the root joint, so predicting the trajectory is important for achieving natural-looking animations. With this system, when the person concerned gives the desired movement to a character or an object in writing, the character can move and the text can be translated into motion. This situation shows that character movement can be made easily with only sentences in the field of animation and game design. With the current developments in the field of artificial intelligence, it has become possible to visualize and mobilize data using only writing. When GPT3 Dall-e (data visualization) and JL2P (data animation) are evaluated from a future perspective away from catastrophic scenarios, it shows that artificial intelligence can create relatively meaningful visuals by trying to establish a correct relationship between text and visual. In the field of mobilization, it can be said that artificial intelligence is

successful at a certain level and the studies in the field continue to develop. Data visualization will allow designers to make quick sketches at the idea stage. It will also allow them to make a quick design plan by putting their existing ideas on the storyboard. Data animation will allow the planned movement to be seen on the character simultaneously. It will positively affect the workforce and time usage. The illustrations created by artificial intelligence will inspire designers and allow them to make creative and original works. Extraordinary experiments on drawing, color, and composition can be made with a single sentence. As one of the design programs used today, it will provide faster results in projects. Artificial intelligence is rapidly developing itself and begins to produce more meaningful results. Visualizing/animating data by using text will become a design tool that anyone who produces in the field of graphics, animation, game design can use soon. This article aims to convey the current developments that artificial intelligence has brought to the data visualization and motion system with an interdisciplinary perspective.