

2024. ÁPRILIS 24.

Az mesterséges intelligencia (MI) hatása a látványtervezésre

Molnár Tamás
R1FYI5
TMDK dolgozat

Pécsi Tudományegyetem
Műszaki és Informatikai Kar
Építészeti tanszék

Az mesterséges intelligencia (MI) hatása a látványtervezésre

Készítette: Molnár Tamás
Évfolyam: 4. évfolyam

Témavezető: Dr. Rák Olivér
Beosztás: adjunktus

Tartalomjegyzék

1. Bevezetés	
1.1. Téma indoklás.....	4.
1.2. A MI ellenség vagy segítség az életünk során.....	4.
1.3. Mi az a Mesterséges Intelligencia (MI) és hogyan működik?.....	5.
2. MI lehetőségek bemutatása	
2.1. A ChatGPT és az emberi agy közötti kapcsolat.....	6.
2.2. Különbség a ChatGPT 3.5 verziója és a ChatGPT 4.0 verziója között.....	7.
2.3. A természetes nyelvi feldolgozás (NLP), nagy nyelvi modell (LLM), Generatív AI.....	9.
3. Kérdőív adatok bemutatása és annak kiértékelése.....	10.
4. Látványtervezés hagyományos lehetőségeinek bemutatása	
4.1. Látványtervezés a digitalizáció előtt.....	13.
4.2. Látványtervezés a digitalizáció után.....	14.
5. Látványtervezés MI-vel támogatott lehetőségei	
5.1. A prompt és a prompt engineering avagy miért fontos a jó prompt írása?.....	15.
5.2. Programok, platformok melyek MI támogatottak.....	18.
5.2.1. MI a AEC: Tervezés alkotásban.....	18.
6. Általam tesztelt eljárások, módszertanok	
6.1. Mely weboldalak/alkalmazások használják már az MI képgenerálást?.....	20.
6.2. PromeAI.....	20.
6.2.1. Promeai főbb jellemzői.....	22.
6.2.2. Hibás képek generálás avagy rossz prompt megadása.....	22.
7. Konklúzió.....	24.
Köszönetnyilvánítás.....	25.
Irodalomjegyzék.....	26.
Hallgatói nyilatkozat.....	28.
Melléklet.....	29.

1. Bevezetés

A kutatásom témájára azért esett a választás, mert én magam is belsőépítészekkel és látványtervezőkkel dolgozom együtt. Így hamar megtapasztaltam azt, hogy milyen is a szakma. Az emberek rögtön és azonnal szeretnék a végeredményt látni. Ebben a szakmában viszont ezt nem lehet kivitelezni, hiszen nagyon sok mindenből tevődik össze, hogy egy adott ház felépüljön, de még ahhoz is sok minden kell, hogy egy egyszerűnek tűnő látványterv elkészüljön. Nem egyszer kaptam már meg én is azt a mondatot, hogy “hát te csak vonalakat húzogatsz egész nap a számítógép előtt nem?”

Amióta jelen van a MI az életünkbe, egész pontosan amióta jobban elterjedt. Hiszen a mesterséges intelligencia már az 1960-as évek óta jelen van. Azóta szinte nap mint nap látjuk, hogy újabb ötletek és újabb oldalak bukkannak fel és próbáljuk ki őket, hogy mire is képes. Köztük van az a lehetőség is amit én dolgozok fel, hogy milyen módon lehet azt elérni, hogy egy fehér modellből egy fotó realisztikus képet generáljunk pár prompt és beállítás segítségével. Nem a saját munkám ellen szeretnék kutatni, hanem megszerettem volna azt vizsgálni, hogy miként tudom segíteni ezzel a munkámat és az ügyfél és tervező közötti kapcsolatot.

Ha neki már annyi is elég, hogy megnézzé hogyan is fog kinézni modelltérbe viszont arra már esetleg nem jut keret, hogy textúrázást, fénybeállítást és minden olyan dolgot kifizessen, ami még a velejárója, akkor miért ne oldhatnánk meg a MI segítségével? Így az ügyfél is elégedett és a tervező is hiszen tudtunk neki segíteni, hogy egy olyan térben élhessen, ami neki otthonos és komfortos. Illetve szerettem volna jobban elsajátítani a MI adta lehetőséget, amellyel fel tudok szállni úgymond a gyorsvonatra és nem leszek lemaradva. Hiszen, ha most nem kezdünk ezzel a témával foglalkozni csak akkor amikor már oda kerülünk, hogy elvárás lesz sajnos akkor már késő lesz, mert annyira előrehaladott lesz a technológia, hogy szinte lehetetlen lesz utolérni magunkat.

A kutatásom során A ChatGPT 4.0 verzióját és a PromAI programot teszteltem és használtam.

1.2. A MI segítség vagy ellenség az életünkbe?

A téma kapcsán sokakban felmerült ez a kérdés és sok vita is folyik ez ügyben hogy a MI elveszi a munkánkat. Ezt sajnos cáfolni nem tudom, hiszen közel 300 millió állás van veszélyben. Ugyan elvesz állásokat, de ezalatt újakat is hoz létre. “A Forbes közölte, hogy a MIT és Bostoni Egyetem jelentése szerint a MI 2025-ig 2.000.000 gyártási dolgozót vált ki.”[1] (saját fordítás)

“A McKinsey Global Institute tanulmánya szerint 2030-ra világszerete a munkavállalók legalább 14%-ának lehet szüksége karrierváltásra a digitalizáció, a robotika és a MI fejlődése miatt.”[1] (saját fordítás) Egy 2023-as adat szerint jelenleg 8 019 876 189 ember él a Földön tehát 1 122 782 667 embernek kell karriertváltania.

A mesterséges intelligencia természetesen lehet ellenség, de lehet segítség is ez attól függ ki hogyan használja. Segítségként számos területen képes javítani és megoldani olyan problémákat amelyeket emberi képességekkel nehéz vagy lehetetlen megoldani.

Mint pl. nagy adathalmazok elemzése, bonyolult számítási feladatok, összetett tervezési feladatok, nyelvek fordítása, extrém környezetek felfedezése. Azonban fontos megjegyezni, hogy ma még csak előny, hogy tudjuk kezelni és használni a MI-t de pár hónap vagy pár éven

belül viszont már elvárás is lehet a cégeknél. Ezt úgy kell elképzelni, hogy ezentúl nem csak az lesz a feltétel, hogy alapszintű excel használata, hanem alapvető ChatGPT 4.0 használata. MI nem fogja kiváltani a diploma megszerzését, mivel az alapvető ismeretek csak még jobbak lesznek, nekünk innentől kezdve nem az lesz a feladatuk, hogy e-mailt, szerződést írunk, hanem ezeket utasításként adjuk be neki és mi ellenőrizzük.

1.3. Mi az a mesterséges intelligencia (MI) és hogyan működik?

A mesterséges intelligenciát más néven MI, vagy az AI-nak nevezik, az angol „Artificial Intelligence” rövidítésből, amely óriási adatbázisból képes automatikusan mintázatokból tanulni azaz felismer bizonyos helyzeteket. Azokból képes komplex cselekvéseket létrehozni. Illetve képesek tanulni a hibáikból, és így növelik a pontosságuk. A MI nem csupán egyetlen technológia, sokkal inkább egy összetett rendszer, ami különböző technikákat ötvözi, mint a gépi tanulás (machine learning), mély tanulás (deep learning), természetes nyelvfeldolgozás (NLP). Valójában mi emberek tanítjuk az MI-t azzal, hogy kérdezzük tőle.

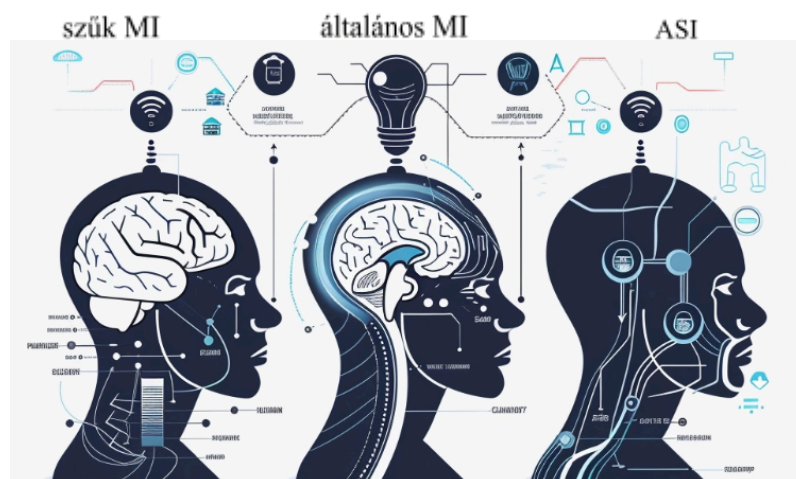
Az MI szó önmagában nagyon félrevezető, hiszen amióta jobban jelen van az életünkbe mindent MI-nek nevezünk. Pedig ezeket bizonyos jelzőkkel kellene ellátnunk, hogy értelmet nyerjen. Ennek a jelzőnek három fajtáját ismerhetjük meg.

Az első ilyen jelző a “amelyet „gyenge MI-nak” is neveznek, a számítógépes rendszer azon képességét jelenti, hogy az embernél hatékonyabban el tud végezni egy szűken meghatározott feladatot.” Viszont ebből a “MI „gondolkodásából” hiányzik a tudatosság és az érzelem”. [2]

A második ilyen jelző “Általános MI, melyet olykor „erős MI-nak” is neveznek, a számítógép azon képességét jelenti, amellyel képes meghaladni az emberi eredményeket bármilyen intellektuális feladatban. Ilyen típusú MI láthatunk olyan filmekben, ahol a robotoknak tudatos gondolatai vannak, és saját céljaiknak megfelelően cselekszenek.”

“Ez a fajta MI már képes lenne nagyon bonyolult problémák megoldására, döntéshozatalra bizonytalan helyzetekben és szélesebb körű feladatokat oldanának meg, mint jelenlegi, szűkebb körű feladatokra betanított MI rendszer.” [2]

A harmadik ilyen jelző “mesterséges szuperintelligencia (ASI) amely egy szuperintelligenciával rendelkező számítógép képes lenne az embert csaknem minden területen túlszárnyalni, többek között a tudományos kreativitásban, az általános bölcsességében és a társadalmi készségekben is”. Viszont az fontos megemlíteni, hogy jelenleg még csak a „erős MI” van jelen a mindennapjainkban. [2]



1. ábra A mesterséges intelligencia 3 típusa

2.1. A ChatGPT és az emberi agy közötti kapcsolat

A ChatGPT megjelenése kreatív forradalmat fog elindítani hacsak már el nem indított.

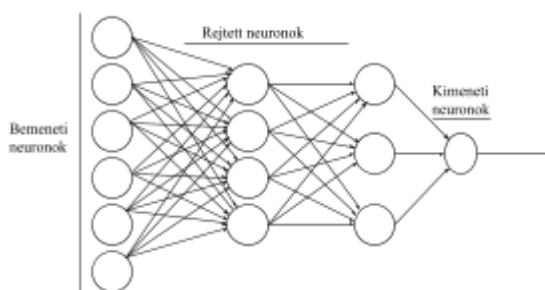
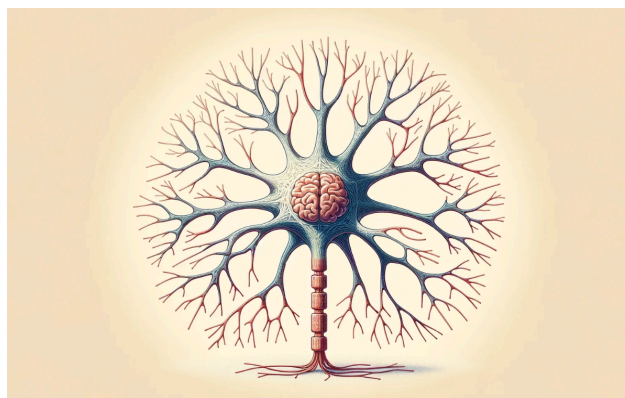
Maga a neurális hálózat ami igazából alkotja ezt az egész rendszert amit mi is létre tudunk hozni, ez egészen hasonlít egy emberi agy idegsejtjeihez.

Az emberi agynak a paraméterszáma 100 billió nagyságrendű és csupán csak 3 GB adatot tud tárolni egész életünk során. 2022. november 30-án megjelent a ChatGPT 3 ennek a paraméterszáma csak 175 milliárd volt így az emberi agy paraméterszámát nem érte el, de így is jóval több adatot és szöveget tudott tárolni egészen pontosan 570 GB-ra volt képes. Ez egy 6 éves folyamatnak köszönhető.

Majd 2023 májusában jött egy újabb mérföldkő a ChatGPT 4-el amely ugyan megegyezik az emberi agy paraméterszámával, de ő már 2-6 TB közötti adatot képes tárolni.

Olyan, mintha mi emberek 138 000 évig olvastunk volna. Ezért képes 5 másodperc alatt összefüggő szöveget létrehozni és válaszolni helyesen vagy részben helyesen adott kérdésekre netán kifejtetni egy adott témát, összehasonlítani különböző építészeti anyagokat, műveket.

Az igazi áttörés még csak most fog jönni a Gemini amely igaz nem éri el az emberi agy paraméterszámát, de jóval több adatot tud tárolni majd ez számügyileg 86 TB szöveges adatot jelent.



2. ábra Az emberi agy és a neurális hálózat közötti hasonlóság

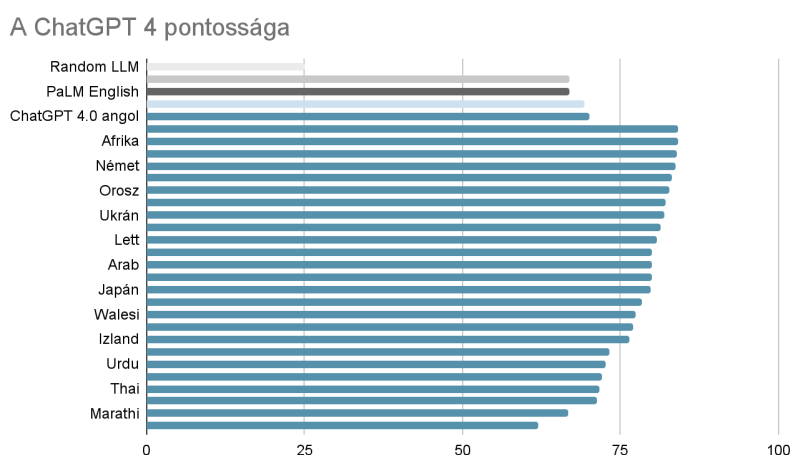
	Mikor készült?	Paraméter száma	Tréningelt adat
Az emberi agy	ie. 300 000	100 billió	<3 GB
ChatGPT - 1	2018	117 millió	4,5 GB szöveges adat
ChatGPT - 2	2019	1,5 milliárd	45 GB szöveges adat
ChatGPT - 3	2020	175 milliárd	570 GB szöveges adat
ChatGPT - 4	2023	1,75-100 billió	2-6 TB szöveges adat
Gemini	2023	30-65 billió	85 TB szöveges adat

3. ábra A ChatGPT fejlődése (Forrás: Amazingai)

2.2. Különbség a ChatGPT 3.5 verziója és a ChatGPT 4.0 verziója között?

“A ChatGPT 3.5 verziója gyorsan felkeltette a mesterséges intelligencia közössége figyelmét az alap a képességével, hogy képes megérteni és létrehozni emberszerű szöveget. Ennek a modellnek azonban bármilyen lenyűgöző is volt, megvoltak a korlátai, különösen a kiterjedt kontextuális információk kezelésében és a többnyelvű kommunikációban.”[3]

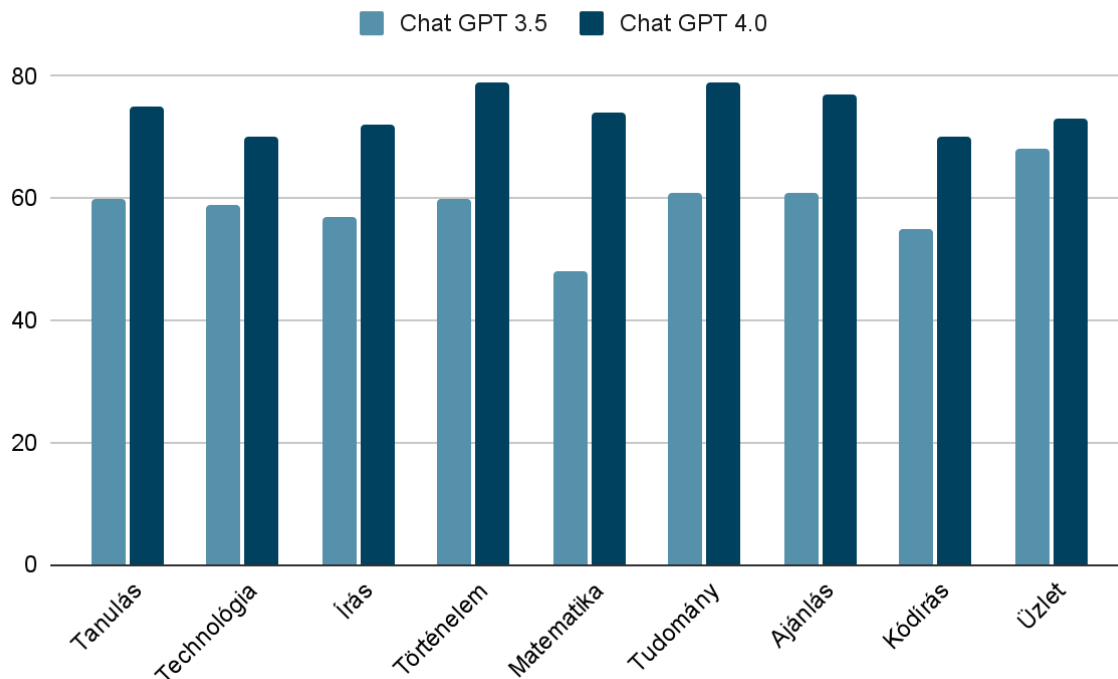
Hiszen a ChatGPT 4.0 verziója sokkal több nyelven tud már kommunikálni számolt be róla az Open AI oldala. Míg a 3.5 verziója csak “25 nyelven beszél beleértve a leggyakrabban használt nyelveket, mint az angol, spanyol, francia, német, kínai, japán, orosz. Addig a 4.0-ás verziója több mint 90 nyelven kommunikál beleértve a legritkább nyelveket is, mint a baszk, katalán, vietnámi, litván, izlandi” [3]



4. ábra GPT-4 3 felvételes pontossága az MMLU-ban több nyelven

“Egy kötetlen beszélgetés során a kettő közötti különbség alig észrevehető, de ha jobban belemerülünk egy adott témában akkor viszont már feltűnnek azok az apró különbségek. A ChatGPT 4.0 sokkal megbízhatóbb, kreatívabb és sokkal árnyaltabb utasításokat is tud már kezelni, mint az elődje a ChatGPT 3.5 verziója.” [3]

Még a ChatGPT 3.5 verziója csak szöveges adatokat tudott létrehozni és csak szöveggel tud válaszolni addig a ChatGPT 4.0 már a DALL-E segítségével már képes képeket generálni, promptokat írni, felismerni adott stílusirányzatokat és képes megmondani, hogy milyen promptot írjak be midjourneybe vagy más képgeneráló szoftverbe ezekekről a későbbiekben ejtek szót. Hiába nagyon nagy a tudástára a ChatGPT 4.0 verziójának attól függetlenül még nem annyira megbízható, mint azt mi hinnénk. Ugyanúgy körültekintőnek kell lenni és utána kell nézni az adott információknak, hogy azok mennyire valósak. “Bár továbbra is valódi probléma, a GPT-4.0 jelentősen csökkent a hallucinációkat a korábbi modellekhez képest. A GPT-4.0 40%-kal magasabb pontszámot ért el, mint a legutóbbi GPT-3.5 a belső kontradiktórus tényértékeléseink során.”[3]



5. ábra Belső tényértékelés kategóriánként (forrás: OpenAI)

Fontos megemlíteni azt, hogy a két verzió között egy igen lényeges különbség áll fenn ami sokak számára elgondolkodtató az az, hogy még a ChatGPT 3.5 verziója ingyenes és bárki számára elérhető korlátlanul tud vele beszélgetni egy vagy több témáról. Addig a 4.-es verziója már fizetős ami 20 dollár havonta ami a mai árfolyammal számolva durván 10.000 Ft kiadást jelent havonta és még a mai napig korlátozva van a beszélgetés nem lehet vele olyan sokáig beszélgetni, mert egy idő után leblokkol pár órára.

Mind a két verziója képes vissza-vissza utalni egy korábbi beszélgetésre és azt folytatni tovább.

2.3. A természetes nyelvi feldolgozás (NLP), nagy nyelvi modell (LLM), Generatív AI

“Az NLP egy módszer, amelyet a számítógépes programok használhatnak az emberi nyelv értelmezésére. Az NLP tulajdonképpen a MI-nak az egyik típusa melyek többnyire gépi tanulásra épülnek fel. Amikor tanítják az NLP akkor különböző szöveggörnyezetbe példákat kapnak szavakra és kifejezésekre, azok értelmezésével együtt.”[4] Tehát ezért nem érti meg néha amikor azt mondom neki, hogy “készíts nekem egy zöld házat”. Neki a zöld ház az azt jelenti, hogy ökológiailag legyen zöld ház. Közben meg én arra gondoltam, hogy legyen zöld színű ház. De miután több ezer példát mutatok meg neki, hogy az én értelmezésem alatt mit jelent a zöld ház azután már ő is azokat a képeket fogja nekem generálni.



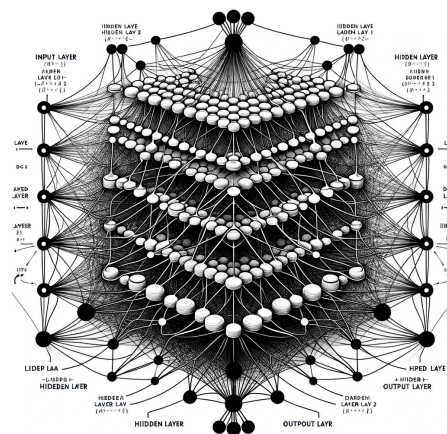
6. ábra Az NLP összetettsége

A **nagy nyelvi modell (LLM)** egyfajta MI program, amely képes felismerni és szöveget generálni többek között más feladatok mellett. “A “nagy” elnevezés nem véletlen, hiszen ezeket egy hatalmas adatbázisra képzik ki. Az LLM egy gépi tanulásra pontosabban egy neurális hálózatra épülnek fel.”[4] Az egyik legismertebb felhasználási területük a generatív MI azaz amikor egy felszólítást utasítást kapnak vagy kérdéseinkre akár kérdéssel válaszol vissza, hogy még pontosabb választ tudjon adni. A legújabb ChatGPT már konkrét esszéket, verseket, képes nekünk írni egy helyes prompt megadása után.



7. ábra Az LLM bonyolult idegi kapcsolattal

“A **generatív MI** a mély tanulásnak egy olyan típusa, amely a szövegek és a képeken felül már képes kódokat és audiovizuális tartalmakat generálni. Ezek is egy nagy adathalmazból tanulnak, hogy képesek legyenek releváns válaszokat adni különféle kérésekre. Gyorsan és hatékonyan képesek kreatív tartalmakat létrehozni, ami vetekedhet egy ember által készített művel is.”[4] A vitatott kérdés ebben, hogy lehet-e ezeket a tartalmakat “újnak” vagy “eredetinek” mondani? A modelleket már képesek vagyunk mi magunk is előállítani, hogyha értünk egy picit a programozás nyelvén. Tehát mi magunk is betudjuk tanítani ezt a hálózatot, hogyha azt mondom neki, hogy “kérlek textúrázd fel a fehér modellem a fekete nappali csoportból”.



8. ábra Generatív MI felépítés

Hiszen abba a csoportba csak fekete nappali képeket és információkat tettem bele, így ő nem fog bele keverni más stílus irányzatot vagy más színharmóniát.

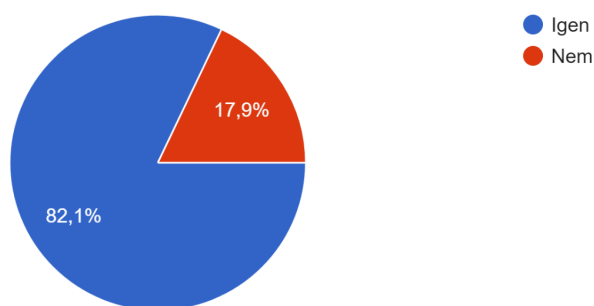
3. Kérdőív adatok bemutatása és annak kiértékelése

A kutatásom során elkészítettem egy kérdőívet melyben felmértem, hogy mennyien ismerik a mesterséges intelligenciát és mire használják azt. Ezt a kérdőívet elkészítettem magyarul és angolul is, melyet megosztottam különböző platformon, de nagy százalékban az egyetemen belül kaptam rá válaszokat.

A kérdőív kiértékelése után választ kaphatunk olyan kérdésekre, hogy hallottak-e már a mesterséges intelligenciáról, használja-e segítségül ezt a lehetőséget a tervezése során vagy milyen nehézségben ütközött már. A MI 67 emberből 12-en nem ismerték.

Hallott-e a MI (Mesterséges Inteligencia) alapú tervezésről?

67 válasz

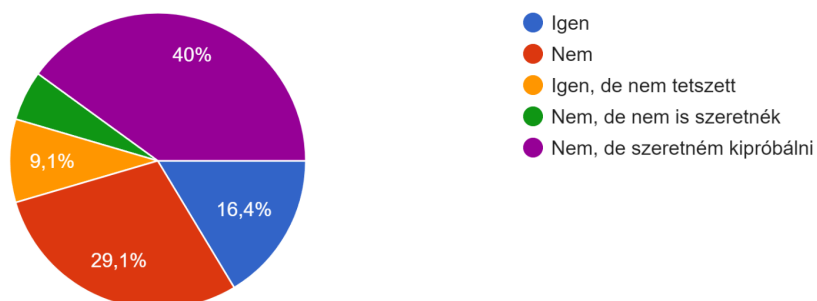


9. ábra kérdőív kiértékelés, Hallott-e már a mesterséges intelligenciáról?

Kíváncsi voltam arra, hogy azok akik eddig nem használták szeretnék-e jobban megismerni vagy beépíteni a minden napjaiba és erre igen nagy százalékban kaptam pozitív választ.

Ön dolgozott-e vagy dolgozik-e MI alapú tervezési eszközökkel az építészeti munkája során?

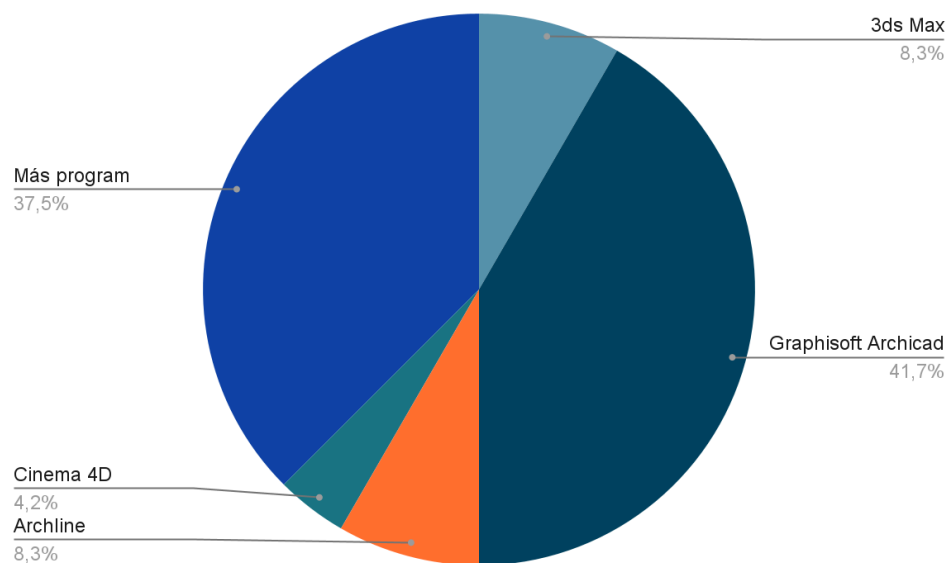
55 válasz



10. ábra kérdőív kiértékelés, Szeretné-e kipróbálni a közeljövőben?

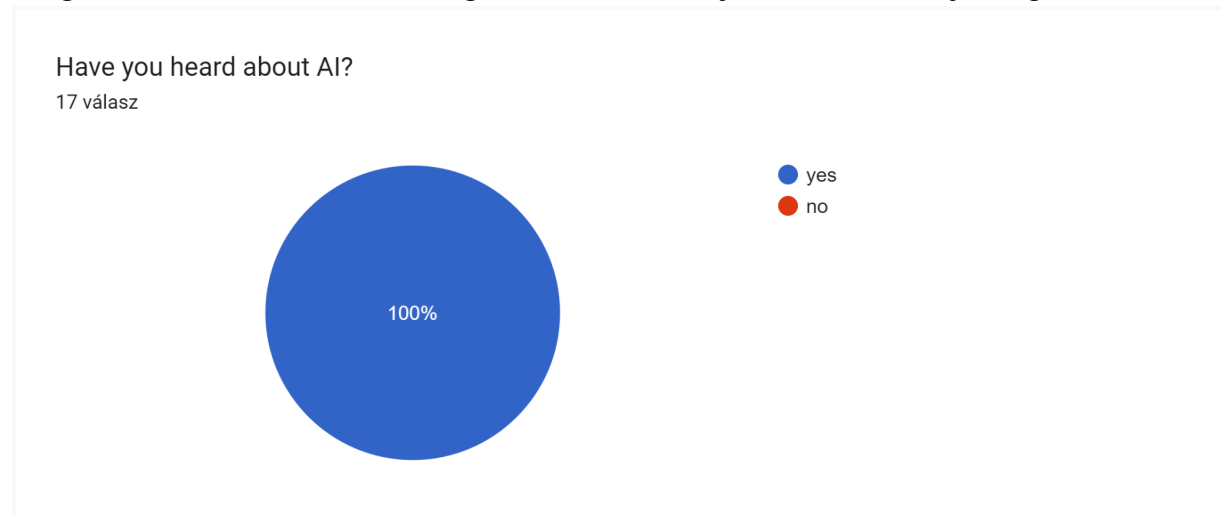
Miközben olvastam a válaszokat arra figyeltem fel, hogy az esetek nagy többségében szinte csak belsőépítészeti látványtervezés foglalkoztak. Természetesen voltak olyanok is, akik pl. homlokzat tervezésre használták, konkrét műszaki leírást irattatnak vele, de valaki AI keresőt használ arra, hogy megtalálja az adott bútort a belső térbe. Viszont az mellett ne menjünk el, hogy ezek még nem tökéletesek, így kíváncsi voltam rá, hogy aki használja ő milyen problémákba ütközött a tesztelés/használata közben. A válasz nagy része az volt, hogy még “elég pontatlan” pl, egy belső, külső tér generálása során, “Egyszerűen nem követi a modellt és torzítja a tereket”. “Nem a kívánt épület alakul ki”, “A nem pontos felvetítés/megtervezése a terveknek.” Ezekből a válaszokból pedig az következtethető ki, hogy nem ásták bele eléggé magukat ezáltal nem úgy használták azt az adott programot ahogyan azt kell. Pontosabban nem megfelelő promptot, nem megfelelő képletet használtak ahhoz, hogy azt lássák amit szeretnének. A dolgozatom további részében bemutatom, hogy milyen oldalak képesek a MI-át kezelni és azokat hogyan kell helyesen irányítani, hogy a várt végeredményt elérjük.

Azért, hogy belelássak abban, hogy ki melyik programot használja ahhoz, hogy tudja segítségül hívni a MI, megkérdeztem azt, hogy “melyik programban dolgozik/dolgozott” Melyből az az eredmény jött, ki hogy az emberek több százalékban az Graphisoft Archicadet használják, de mellette a ChatGPT valamelyik verzióját is akárcsak keresésre, de van hogy műszaki leírás pontokban szedésére. Vannak programok amelyeket még használnak ilyen a Cinema 4D, az Adobe Photoshop MI-vel támogatott verziója, Procreate.

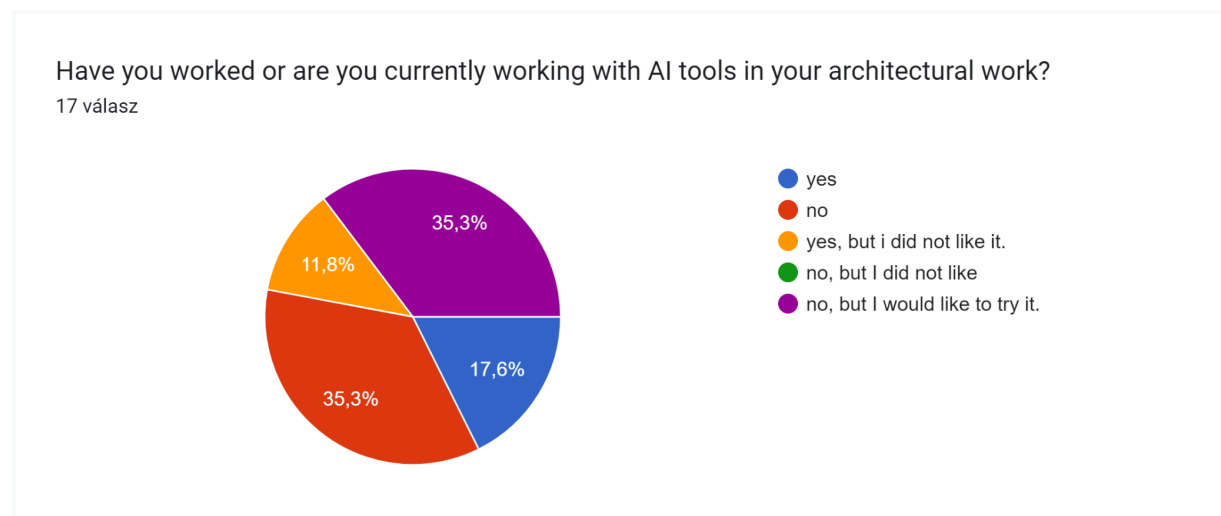


11. ábra kérdőív kiértékelés, MI segített tervezés

Ahogy korábban említettem, a tesztet angolul is elkészítettem, hogy összehasonlítsam a különböző eredményeket. Ebben az esetben csak az egyetem külföldi hallgatóitól kértem segítséget. A kérdések mindkét nyelvi változatban azonosak voltak, azonban angolul csak 17 hallgató válaszolt. Az első kérdés eredményei különösen meglepőek voltak, hiszen a külföldi hallgatók körében az MI ismerettségéről szóló kérdés teljes, 100%-os elterjedtséget mutattak.



12. ábra külföldi kérdőív kiértékelés, Halott-e már a mesterséges intelligenciáról?



13. ábra külföldi kérdőív kiértékelés, Szeretné-e kipróbálni a közeljövőben?

4. A látványtervezés hagyományos lehetőségeinek bemutatása.

4.1. A látványtervezés a digitalizáció előtt

A Britannica cikke szerint “a grafikai tervezés területén a digitális forradalom előtt a tervezők kézzel rajzoltak, hogy vizualizálják a koncepciókat; a típust kézzel rendelték meg és a képekkel együtt fizikai alapon állították össze a terveket a nyomtatás előkészítése érdekében és a képek típuspróbáját és fotostatáját nehéz papírra vagy táblára állítottuk össze fényképészeti sokszorosítás és lemezkészítés céljából.”[5] (saját fordítás)

A látványtervezés digitális előtti módszerei jelentős változásokon mentek keresztül az évek során. Az ArchDaily cikkéből kiderül, hogy “az építészetben a vizuális képi technikák, mint a kézi rajzolás és a lineáris perspektíva alkalmazása, évszázadokon át voltak a tervezők fő eszközei. Ezek a módszereket tették az építészek számára, hogy részletes terveket készítsenek, még a tényleges építkezés megkezdődött volna.” [6] (saját fordítás)

A látványtervezés módszerei közé tartoznak azok a technikák, amelyek még a digitális eszközök elterjedése előtt alkalmaztak a tervek vagy a koncepciók vizuális megjelenítésére.

Ezt a látványtervezési lehetőséget a napjainkban már nagyon kevés tervező, lakberendező, tájépítész használja inkább már a digitális módszert használjuk. Ami teljesen érthető, hiszen jóval könnyebb módosítani egy digitális látványterven, mint egy kézzel rajzolt látványon. Nézzünk pár hagyományosan elkészített látványterv lehetőséget.

Az első ilyen ami szinte az egyik legelterjedtebb az a rajzolás és festészet. Ennek is számos lehetőségei. Hiszen lehet, akvarellal, temperával, vízfestékkel, grafitral, tűfilccel.

A másik ilyen lehetőség a makettezés ami még kedvezőbb, mint egy papírra vetett 3D-s rajz hiszen itt az ügyfél háromdimenziós formában is megtudja vizsgálni az adott épületet.



15. ábra Általam készített hagyományos látványterv akvarell + diópáccal



16. ábra Általam készített hagyományos látványterv ecsettel + skiccpausz

4.2. A látványtervezés a digitalizáció után

“A belsőépítészeti látványtervezés egy olyan háromdimenziós térmodell létrehozásának folyamata, amely a lehető legközelebb áll a valósághoz. A speciális 3D-s modellező és vizualizációs programok használatának köszönhetően az egyszerű 2D-s alaprajzok a legvalósághűbb részleteket tartalmazó 3D-s tervekkel alakulnak. A tervezési render lehetővé teszi az otthon, iroda, bevásárlóközpont, szálloda vagy egyéb helyiség jövőbeli megjelenését a projekt fejlesztésének szakaszában.” [7] (saját fordítás)

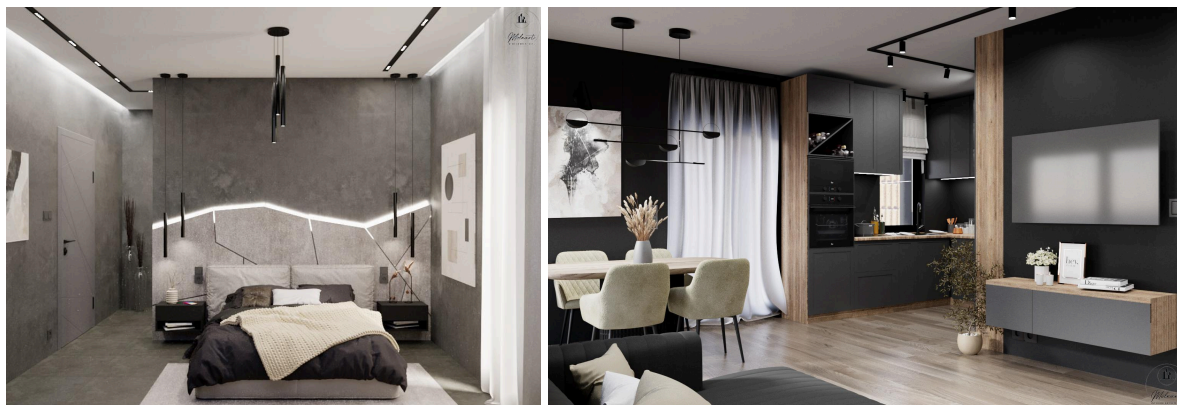
A látványtervezés digitális eszközei forradalmasították a tervezési folyamatokat, lehetővé téve gyorsabb, hatékonyabb és részletesebb vizualizációkat. Bemutatok néhány példát a digitális látványtervezés során.

Ilyenek a számítógépes 3D modellezés, mint pl.: Autodesk Revit, Graphisoft Archicad, SketchUp Desktop melyek lehetővé teszik a nagyom részletes kidolgozást és könnyen módosíthatjuk és nézhetjük meg különböző szemszögből.

Ilyenek a renderelés melyek képesek arra, hogy teljesen fotorealistikus képet tudjon előállítani, mint pl.: V-Ray, Lumion, 3ds Max, Cinema 4d viszont az utóbbi kettőhöz szükség van különböző render szoftver ilyen lehet a corona render.

Természetesen van lehetőségünk 3D animációra is, hogy videó keretében tudjuk megmutatni a megrendelőnek az adott épületet. Erre alkalmas programok lehetnek a Cinema 4D, Blender, Lumion.

A digitális eszközök jelentősen bővítik a tervezők képességét és lehetővé téve azt, hogy az építészek gyorsabb reagáljanak a változásokra, hatékonyabban kommináljanak és ezáltal magasabb színvonalú terveket hozzanak létre. Illetve “a kulcs itt az interneten való láthatóság; minél jobb a munkája minősége, annál több figyelmet von maga után. Pontosabban, a nagy renderelési felbontással továbbfejlesztett 3D-s belső vizualizáció nemcsak a projektbemutatók minőségét emeli, hanem a professzionalizmust is.” [6] (saját fordítás)

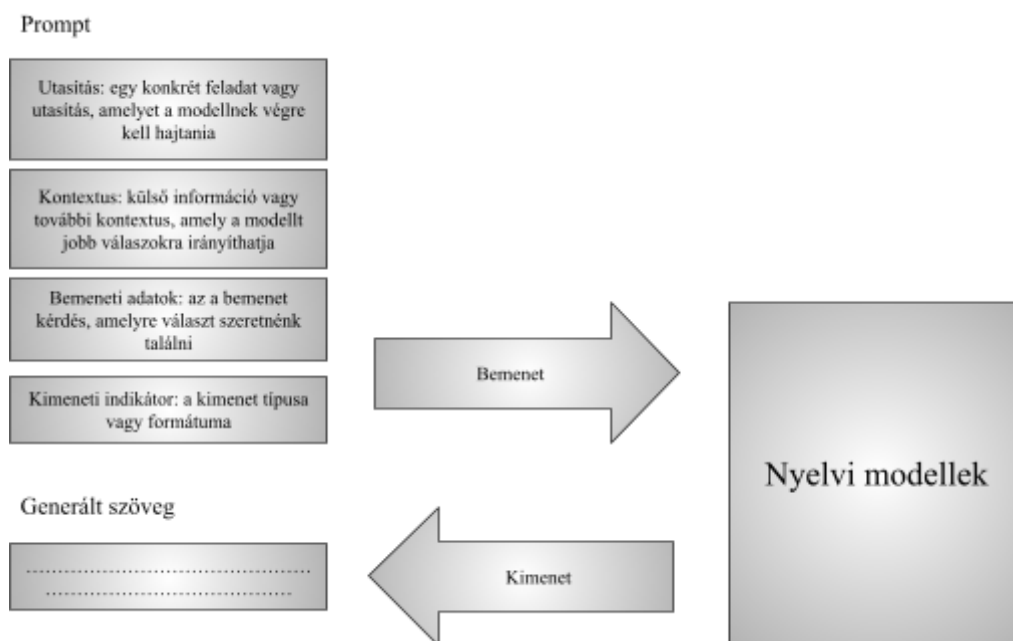


17. ábra Digitális belsőépítészeti látványtervek Cinema 4D + corona render

5. Látványtervezés MI-vel támogatott lehetőségei

5.1. A prompt és a prompt engineering avagy miért fontos a jó prompt írása?

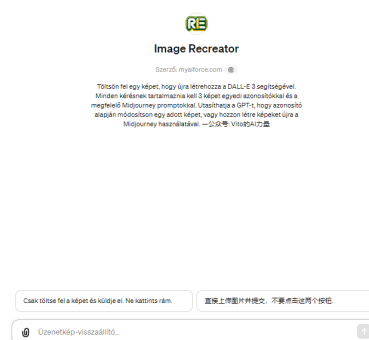
A prompt az az input vagy utasítás, amit a nyelvi modellbe táplálunk annak érdekében, hogy a legjobb választ kapjuk. Minél specifikusabb, árnyaltabb a promptunk, annál jobb megoldást (választ) fogunk kapni a rendszertől. A válasz lehet szöveg, kép, hang, videó vagy akár egy kódsor is. A prompt engineering azt a tervezési folyamatot jelenti, hogy hogyan adjunk utasítást a legjobb módon a nyelvi modellnek, megfelelő kontextust és útmutatást biztosítva. Egyszerűbben fogalmazva: hogyan tegyünk fel jó kérdést/kérést? Minél részletgazdagabb, specifikusabb az input amit betáplálunk, annál jobb lesz a válasz/megoldás amit kapunk.[8]
A jó prompt elemei a következők:



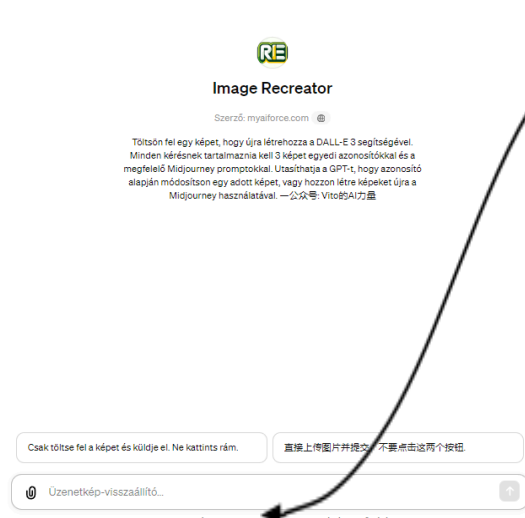
18. ábra Helyes prompt írása (forrás: Amazingai)

A prompt írásnál fontos arra figyelni, hogy nem azonnal kell neki megadni a teljes utasítást. Ahhoz, hogy jól sikerüljön a kép generálása be kell neki mutatni, hogy mi a tervünk a képpel. Mivel ez egy iteratív azaz ismétlődő folyamat ezért sokat kell vele kísérletezni. Kezdjük egyszerűen és folyamatosan adjunk hozzá a további elemeket. Ha nagy feladatról van szó akkor szedjük szét több kisebb részfeladatra. A ChatGPT4.0 verziója már képes arra, hogy egy képről megmondja azt, hogy milyen prompt alapján tudjuk azt a képet generálni. Azaz, ha nekünk tetszik egy kép bármely platformról akkor ha elküldjük neki akkor ő megmondja annak az utasítását amit a midjourney meg tudunk adni.
Képgenerálás lépései:

1. nyissuk meg a <https://chat.openai.com/g/g-ksMYQwohp-image-recreator> oldalt



2. Töltsük fel az inspirált képet ide:



Töltsünk fel egy képet ide!
Egyszerűen csak illeszd
be!
Három variációt kreál,
amihez a DALL-E 3-at
használja

3. A kép, mint inspiráció:



4. ChatGPT 4.0 által generált 3 kép az inspiráció szerint:



5. Minden generált kép alá létrehoz egy egyedi kódot:

 Image Recreator



Image xt: RNeqLvbi5my8Dop

```
sql Copy code  
  
/image prompt: Medium: Photo. Subject: Modern minimalist living room, charcoal sofa
```

Egyedi ID kóddal látja el a képeket: Minden generált képhez saját azonosítót kap, ami megkönnyíti a későbbi finomhangolásokat.

Mijourner promptok: Minden képhez egy Midjourney felszólítást is kapunk

A részletes promptok segítenek a generált képek pontosításában.

ChatGPT 4.0 által generált prompt:

“/image prompt: Medium: Photo. Subject: Modern minimalist living room, charcoal sofa, plush cushion, beige rug, round glass coffee table, black metal frame, lit candle, vertical black slats, gray wall, window with blinds, elegant statuette, dark entertainment unit, white ceiling, recessed lighting, taupe curtain. Emotion: Calm. Lighting: Soft ambient. Scene: Interior design showcase. Style: Modern minimalist --ar 1:1”

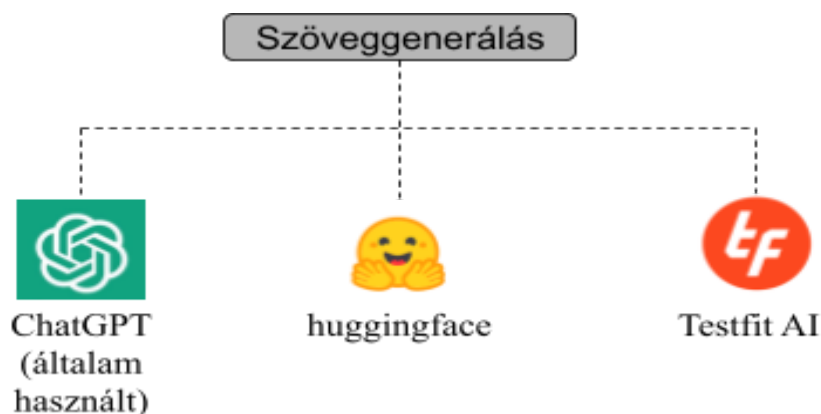
5.2. Programok, platformok melyek MI támogatottak

Vannak már olyan tervező programok amelyekbe már belevan építve eleve a renderelőjébe a MI lehetőség. Ennek az előnye az az, hogy azonnal tudjuk használni és nem kell kimenteni a képet, beilleszteni. Hanem csak rákattintunk és beírjuk azt, hogy pontosan mit is szeretnénk látni. Az egyik ilyen program ami legjobban elterjedt az építészek körében az az Archicad 27-es verziója. Viszont fontos megjegyezni, hogy ez jelenleg a magyar verzióban nem elérhető csak az angol verziójában.

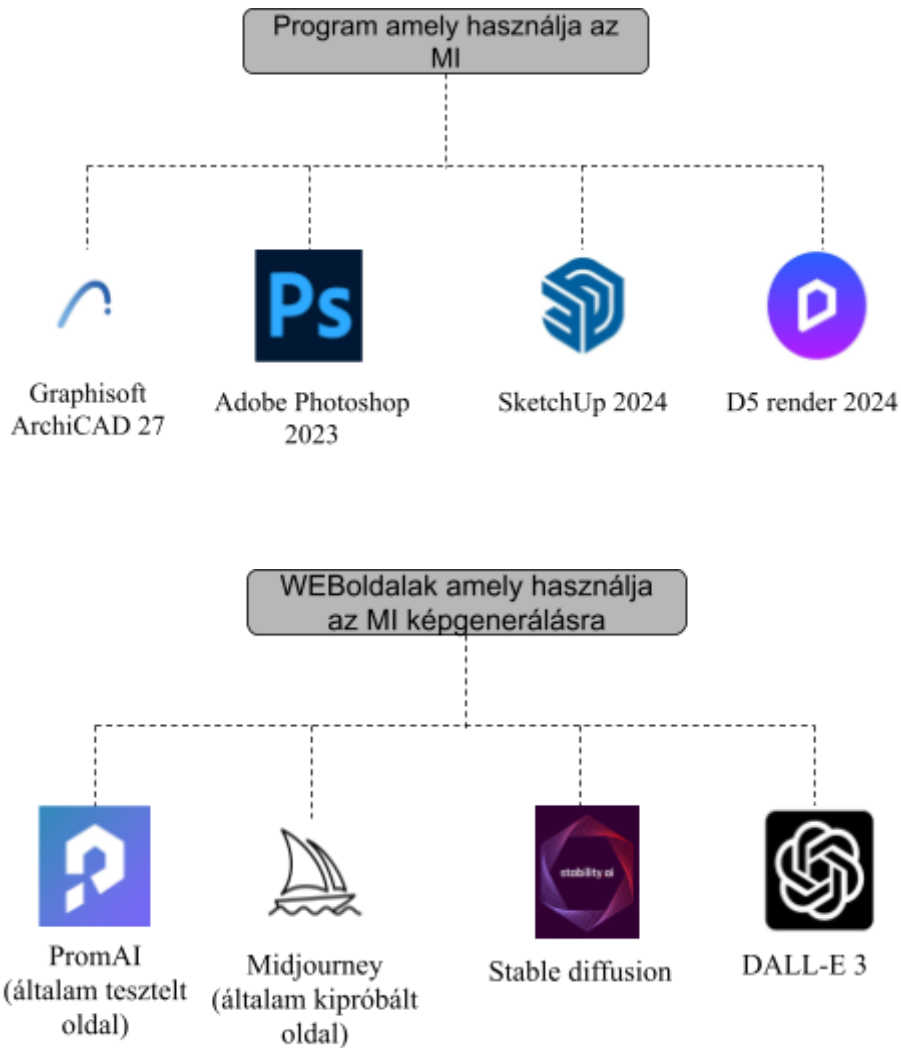
Természetesen nem csak a szerkesztő programok képesek erre, hanem a grafikáért felelős programok is, mint az adobe photoshop 2023-as verziója.

5.2.1.MI a AEC: Tervezés alkotásban

A látványtervezés területén egyre jobban lesz szükség a látványtervezés területén különösen azoknak akik vizuális tartalmakat hoznak létre. Egyes alkalmazások már képesek arra, hogy pillanatok alatt szövegből képet generáljunk ezáltal a koncepció alkotást felgyorsítja. Viszont ahhoz, hogy jó képet kapjunk elengedhetetlen a jó prompt írása amit az esetek nagy többségében a szöveggeneráló alkalmazások képesek rá. A szöveggenerálót nem csak erre lehet használni. Hanem kiválóan működik két anyag összehasonlításában is. A ChatGPT 4.0 verzióját teszteltem.



Az építészet területén egyre több olyan alkalmazás van amiben már a saját rendelőjében vagy magába a programba bele van építve a mesterséges intelligencia segítségként. Ezek leginkább a legújabb programokban találhatók meg amely már 2023-2024-es fejlesztésen átesett át. Az alábbi kép illusztráció csak pár darabot sorol fel.



Van arra is lehetőségünk, hogy ne kelljen újabb programokat, mint például a d5 rednert megtanulunk csak azért, hogy egy szép képet le tudjunk generáltatni vele. Ezért létrejöttek és a mai napig hoznak létre újabb és újabb weboldalakokat, hogy akár felgyorsítsák a munkánkat vagy csak játszadozzunk az adott weboldal-as verzióval. Ezeknek még amellett, hogy viszonylag kevés idő alatt (1-2 nap) meg lehet tanulni a kezelését előnyt jelenthet az is, hogy nem kell letölteni, hiszen bármely eszközön használható. A midjourney, Stable diffusion illetve a DALL-E 3 egyelőre nem képes képet fogadni feladatként, ő ezeket csak mintaként tudja kezelni és ugyanolyan képet hoz létre. Fontos megjegyezni azt, hogy a midjourney egy Discordon keresztül működő nyilvános felület. Ami annyit jelent, hogy mindenki látja. A pro verzióban már megnyílik egy külön chat szoba amit már csak én látok. A DALL-E 3 az A ChatGPT 4.0-ás verziójába van beépítve. Ugyanúgy csak előfizetés után jelenik meg ez a verzió ahogy ezt már említettem a “A prompt és a prompt engineering avagy miért fontos a jó prompt írása?” fejezetben.

6. Általam tesztelt eljárások, módszertanok

6.1. Mely weboldalak/alkalmazások használják már az MI képgenerálást?

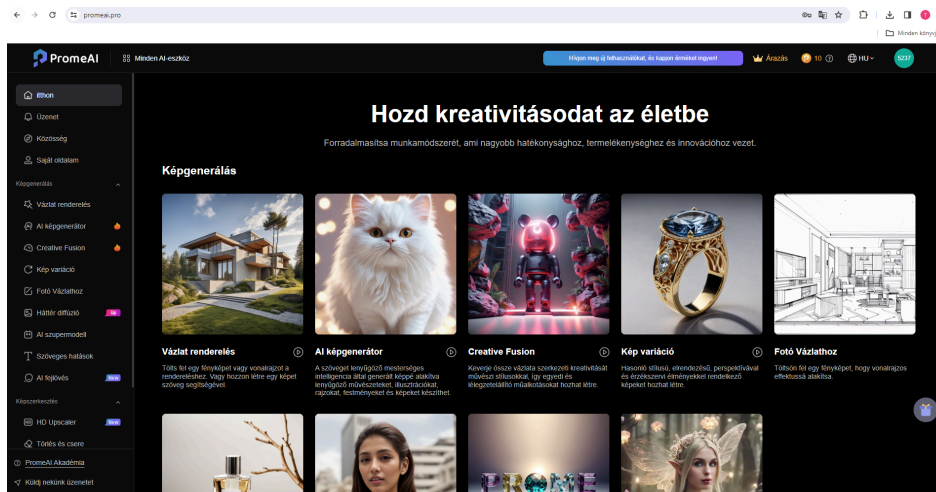
Amióta berobbant az MI azóta hétről hétre hoznak létre újabb platformokat amely pillanatok alatt képes fotorealisztikus képeket generálni csupán pár prompt megadása után.

6.2. Általam fejlesztett módszertan.

A kutatásom során számos alkalommal vizsgáltam meg a témát különböző kísérleti körülmények között és ezáltal széles körű és átfogó eredményekre tettem szert, amelyek alapján mélyebb betekintést nyertem a témában. A változások során különféle paramétereket teszteltem, hogy teljes körű megértését szerezzek, így hozzájárulva az adott tudományterület ismereteinek bővítéséhez.

1. lépés. megmodelleztem egy belső teret, melyekhez nem használtam semmilyen konkrét textúrát. Minden egyes modellre ráhúztam egy sima fehér textúrát (kivéve az üveget állítottam be, hogy eljőjön a kinti napfény)
2. lépés. a belső teret lerendereltem, hogy kapjak egy éles képet amit lementettem Full HD+-ban (1920 x 1080), mindezt JPG formátumban.
3. Ahogy azt már említettem egyelőre a PromAI képes képet fogadni úgy, hogy azzal dolgozzon és textúrázza fel illetve azon végezzen módosítást. Akár

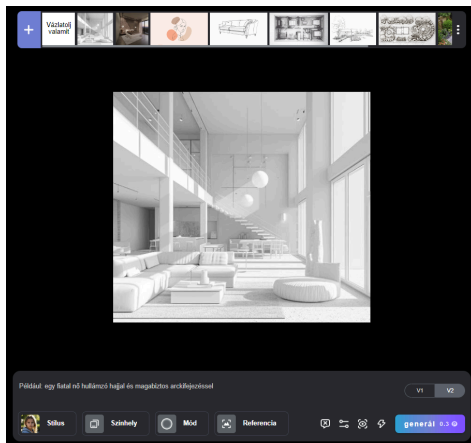
6.2. Promeai: A Promeai-ban két lehetőségünk van az első lehetőség az az, hogy mi beállítunk egy referencia képet ami alapján számolhatjuk ki a képet. A másik lehetőség, hogy



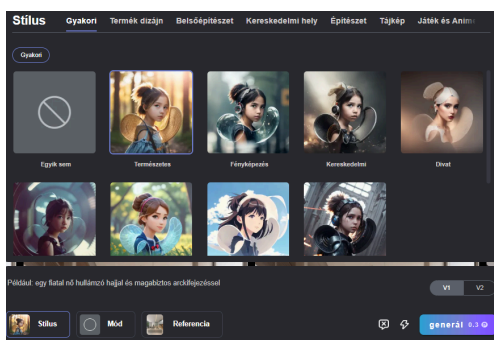
promptot adunk meg neki. Ettől függetlenül ugyanúgy be kell állítani az elsőben említett beállításokat, hogy azt kapjuk amit szeretnénk látni. Ez egy webes felület melyből képből csinál képet vagy akár rajszerű

vázlatot. Ennek az az előnye, hogy nem kell semmilyen alaptudás sem semmilyen promptot használni, csak beállítjuk amit szeretnénk látni. Egy kis ideig ingyenes majd amikor elfogy a kredit rendszerünk fizetőssé válik. Ez pontosan 10 kreditet jelent minden képgenerálás 0,3 kreditet használ fel.

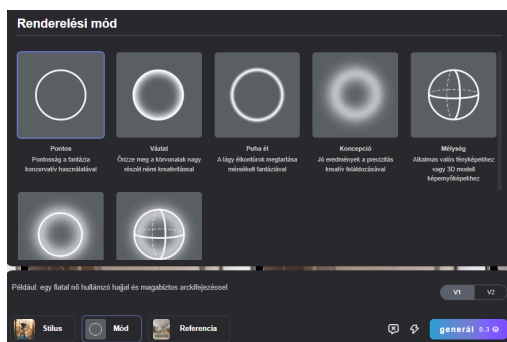
Vázlat renderelés:



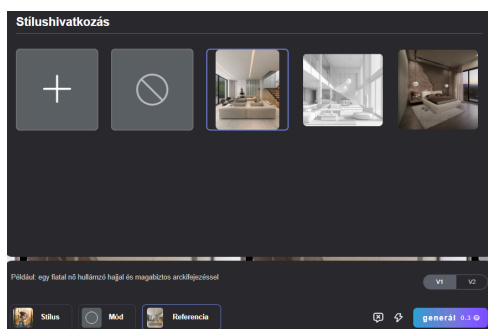
Kattintsunk a + jelre és adjuk meg azt a képet, amiből szeretnénk vázlatot készíteni. Majd kattints át a V1-re.



19. ábra kezelőfelület bemutatása



olyat amire mi nem is gondoltunk azelőtt még nem láttuk a képet



Állítsuk be a stílust, hogy milyen stílusban szeretnénk, hogy készüljön el kép.

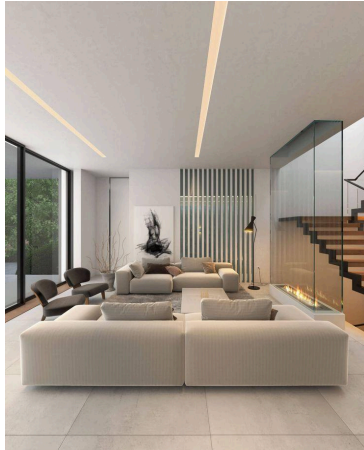
Hiszen a Promai-ban megtalálható rengeteg stílusirányzat azért is fontos, hogy ezt beállítsuk, hiszen mi nem egy mesébe vagy fantázia filmbe szeretnénk megjeleníteni a képeket, hanem egy teljesen reális képet szeretnénk létrehozni

Jelen esetben a kép természetes formában fog

Állítsuk be a renderelési módot.

Az alkalmazás felkínál nekünk 7 fajta renderelési módot a példában az első verziót azaz a pontos tehát teljesen leköveti az adott modellt és nem tesz hozzá semmi plusz kiegészítőt. Ha még nem vagyunk teljesen biztosak abban, hogy teljesen készen van a modell vagy netán üresnek tűnik a kép akkor olyanra kattintsunk amely nem követi le teljesen a modellt ezáltal még hozzá ad valami

Majd adjunk meg referencia képet a + jelre kattintva ami azért szükséges, mert így tudja, hogy én milyen képet szeretnék látni. Így nincs szükség különböző beállításokra. Mivel ezeket a színeket fogja használni vagy közel ugyanezeket.



Referencia kép melyet hozzáadtam.

Amiért erre a képre esett a választás, mint referencia, mert én is hasonló stílusban hasonló színben szeretném látni a nappalimat



Végeredmény

Hátránya: nem lehet plusz kiegészítőket hozzáadni, mint pl “tegyél a sarokba egy virágot” Illetve nem 100%-an követi le a modellt

20. ábra végeredmény bemutatása

6.2.1. PromeAI főbb jellemzői:

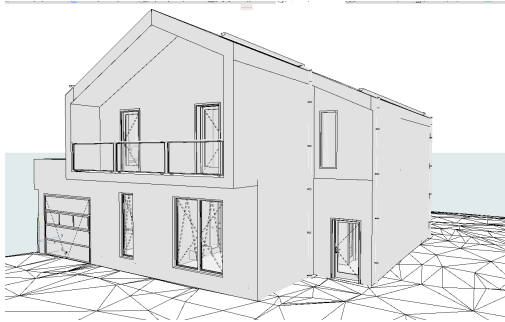
Ami számomra az egyik legfontosabb volt a kutatásomkor, hogy nem kell külön programként kezelni, hiszen ez egy webes felület, így ha van internetünk bárhol és bármilyen eszközről kezelhető volt, legyen szó, telefonról, tabletről, asztali gépről és laptopról. Ami még fontos szempont lehet, hogy a látványterv teljesen valósághűen tudja kezelni a terveket ezáltal egy pontos anyaghasználatot és megvilágítást kapunk végeredményről nem csak egy sematikus szín használatot. A promAI-nak van egy prémium lehetősége amellyel azt érzük el, hogy nagyobb felbontásban tudjuk lementeni a képeket akár 4k-s felbontásban is.

Maga a **PomeaAI-t a DALL-E** nevezetű képgeneráló rendszer működteti melyet az OpenAI fejlesztett ki és üzemeltet a mai napig.

6.2.2. Hibás képek generálás avagy rossz prompt megadása

A promAI-ban van lehetőség arra, hogy ne mi magunk állítsuk be a stílusirányzatot hanem egy egy szóval adjuk meg neki, hogy mit is szeretnék látni. Bemutatom azt, hogy miért is fontos a helyes prompt fogalmazás.

Jelen szituációban Archicadben megmodelleztem a házat és egyszerű képernyőképet készítettem, majd beillesztettem a webes felületre. Promptnak csak annyit írtam, hogy “green house” nem fogalmaztam egyértelműen. A generáló ez alatt az értette, hogy legyen ÖKO házam pedig én azt szerettem volna, hogy legyen zöld a homlokzata. A helyes prompt pedig az, hogy “green facade”



Általam készített modell Graphisoft Archicad 27-es oktatási verzióban



PromAI által generált kép rossz prompt megadása után



PromAI által generált kép jó prompt megadása után

7. Konklúzió

A mesterséges intelligencia jelentős hatást gyakorol a látványtervezés területére, pedig még csak kevesebb, mint egy év telt el és ez várhatóan tovább fog növekedni a jövőben.

A kutatás során megvizsgáltam azt, hogy mely program, weboldal képes arra, hogy csak a fehér modellt vegye alapul és azt textúrázza fel és ne adjon hozzá semmi kiegészítőt. Az elterjedtek közül, mint a Midjourney a Stable diffusion vagy akár ChatGPT 4.0 által vezérelt DALL-E 3 nem alkalmasak erre. Ők csak a promptokból tudnak képet generálni és ezeket tudjuk folyamatosan frissíteni, módosítani, nagyítani plusz kiegészítőket hozzáadni.

Míg a PromAI képes mind erre. Természetesen neki is vannak még hiányosságai. Ez alatt értem azt, hogy ő neki nem lehet utasításként még adni azt, hogy tegyél egyes helyekre kiegészítőket. Itt ennyi utómunka szükséges, hogy ha mindenképpen szeretnénk plusz kiegészítőket tenni az adott látványtervre, de magán a képen nem kell használni utómunkát.

Az alkalmazások ingyenes köre egyre jobban kezd szűkülni a nagy felhasználó kör miatt.

Összességében a MI a “képből képet” területen még nincsen teljesen kifejlesztve. Minden programban vannak hiányosságok, de még jelenleg nincsen egyetlen egy olyan program sem amely mindegyiket tudná egybe.

Ezeket a programokat ha megfelelően tudjuk használni nagyban elősegíti a munkákat és egyben felgyorsítja. Természetesen ehhez meg kell találni az arany középutat.

Illetve használhatjuk akár olyan megoldás javaslatra is amire mi lehet nem is gondoltunk volna ezt akár nevezhetjük úgy, hogy “hiba, mint esztétika” és ez adja meg a külső/ belső tér különlegességét. A további céljaim olyan saját nagy nyelvi modell létrehozása amit már én irányítok. Tehát már az én stílus készletemből válogat és abból öltözteti fel azt a bizonyos fehér modellt.

8. Köszönetnyilvánítás

Szeretném kifejezni a köszönetemet minden egyes személynek aki hozzájárult és támogatott a dolgozatom elkészítésben. Elsősorban köszönöm Dr. Rák Olivér adjunktus Úrnak, aki nemcsak hogy témavezetőként irányított, de folyamatosan segítette a munkám a kutatásom során.

Hálás vagyok a Pécsi Tudományegyetem Műszaki és Informatikai Kar részéről nyújtott anyagi támogatásért, amely elengedhetetlen volt a projekt létrejöttéhez. Köszönöm a szaktársaimnak, akik segítettek és támogattak a kutatásomban értékes visszajelzésekkel.

Illetve köszönöm szépen a különböző egyetemi tanároknak, Dr. Széll Attila Béla, Dr. Németh Pál szakmai támogatást és útmutatást, amely hozzájárult a sikeres kutatásom és a téma mélyebb megértéséhez.

Végül, de nem utolsó sorban, szeretném megköszönni a családomnak a támogatást, amelyet mindig is nyújtottak nekem hiszen nélkülük nem jutottam volna el idáig.

Irodalomjegyzék:

1. MARK TALMAGE-ROSTRON. (2024. JANUÁR 10). *Nexford University*. Letöltés dátuma: 2024.. MÁRCIUS 13., forrás: Nexford University: <https://www.nexford.edu/insights/how-will-ai-affect-jobs>
2. Microsoft. (dátum nélk.). *Microsoft Azure*. Letöltés dátuma: 2024. Április 09., forrás: <https://azure.microsoft.com/hu-hu/resources/cloud-computing-dictionary/what-is-artificial-intelligence/#%C3%B6nvezet%C5%91-aut%C3%B3k>
3. OpenAI. (2023.. Március 13). *OpenAI*. Letöltés dátuma: 2024.. Március 1, forrás: <https://openai.com/research/gpt-4>
4. Fedorovics, V. (dátum nélk.). *Amazinum*. Letöltés dátuma: 2024.. Február 26., forrás: Amazinum: <https://amazinum.com/insights/what-is-nlp-and-how-it-is-implemented-in-our-lives/>
5. Meggs, P. B. (dátum nélk.). *britannica*. Letöltés dátuma: 2024.. Április 08., forrás: <https://www.britannica.com/art/graphic-design/The-digital-revolution>
6. Baraya, S. (2020.. Július 10.). *Archdaily*. Letöltés dátuma: 2024.. Április 14., forrás: <https://www.archdaily.com/942862/the-evolution-of-visual-representation-in-architecture-and-how-it-will-continue-to-change>
7. Maksymova, I. (2023.. Február 15.). *Applet 3D*. Letöltés dátuma: 2024.. Április 14., forrás: <https://www.archdaily.com/942862/the-evolution-of-visual-representation-in-architecture-and-how-it-will-continue-to-change>
8. Bereczki Nóra. (2023. Július 19). *LABS*. Letöltés dátuma: 2024. Április 2., forrás: LABS: <https://www.xlabs.hu/blog/mi-az-a-prompt-engineering-ertsuk-meg-egy-peldan-keresztul>
9. Taylar, R. (2023.. november 17). *Medium*. Letöltés dátuma: 2024.. Február 22, forrás: <https://medium.com/@RickTaylar/the-only-prompt-formula-youll-ever-need-for-chatgpt-6946599c2912>
10. Charlie Giattino, E. M. (dátum nélk.). *Ourworldindata*. Letöltés dátuma: 2024.. Március 10., forrás: <https://ourworldindata.org/artificial-intelligence#article-licence>

11. Carpo, M. (2018). Letöltés dátuma: 2024.. Március 3, forrás:
<https://mariocarpo.com/articles/excessive-resolution-designers-meet-the-second-coming-of-artificial-intelligence>
12. Ostwald, M. J. (2024.. Január). *ResearchGate*. Letöltés dátuma: 2024. Február 21., forrás:
https://www.researchgate.net/publication/377355618_Multimodality_and_Architectural_Technology
13. Máté, H. (2023.. Február 13.). *Építészforum*. Letöltés dátuma: 2024. Március 19., forrás:
<https://epiteszforum.hu/es-akkor-jottek-a-robotok-es-elvettek-a-munkankat-avagy-a-mesterseges-intelligencia-hatasa-az-epiteszetre>

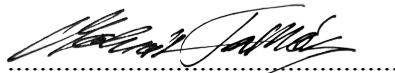
Hallgatói Nyilatkozat

Alulírott hallgató kijelentem, hogy a Tudományos és Művészeti Diákköri konferenciára benyújtott dolgozat a saját munkám eredménye.

A felhasznált szakirodalmat és eszközöket azonosíthatóan közöltem, és szakszerűen kezeltem. Egyéb jelentősebb segítséget nem vettem igénybe.

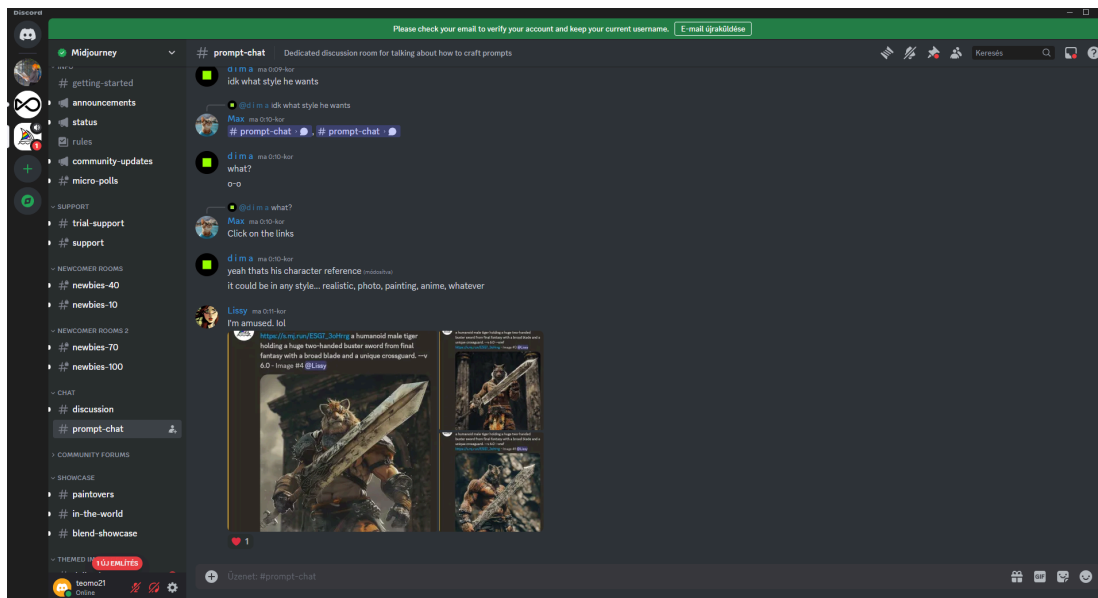
Az elkészült Tudományos és Művészeti Diákköri konferenciára leadott kutató munkám által talált eredményeket a Pécsi Tudományegyetem, Műszaki és Informatikai kara saját céljaira térítés nélkül felhasználhatja.

Pécs, 2024. Április 19.

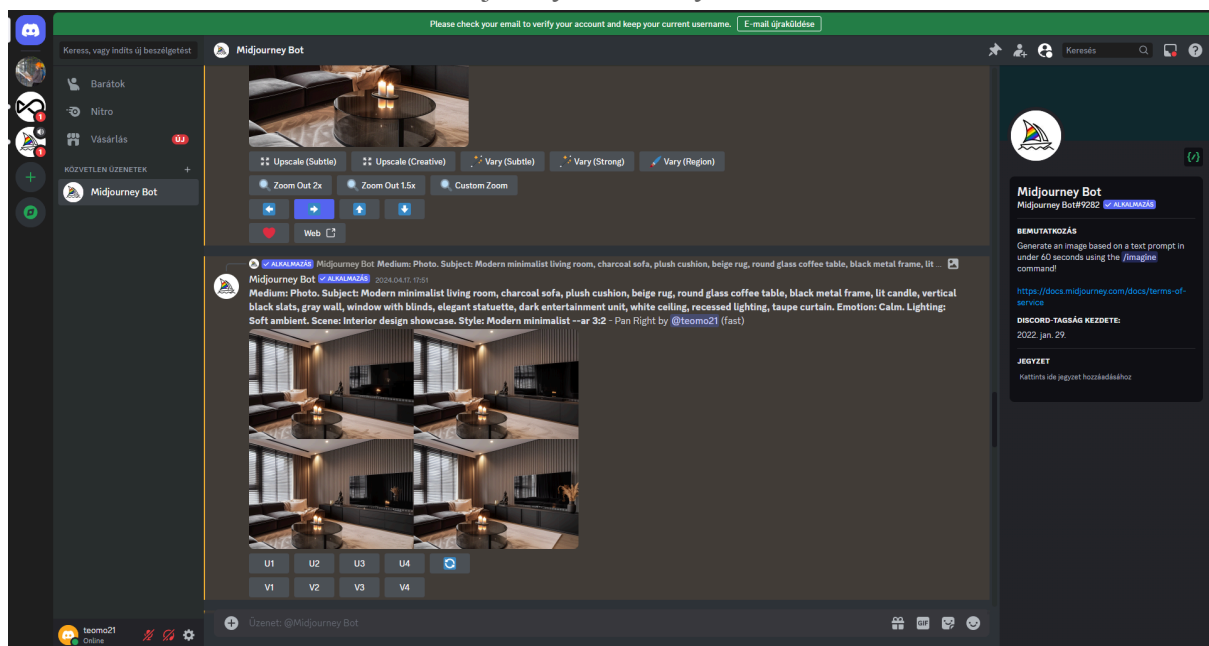


Hallgató aláírása

Mellékletek



21. ábra Midjourney discord nyilvános felület



22. ábra Midjourney discord privát felület



23. ábra rosszul generált képek