

„Kreálok valamit, csak így szavakból!” – mesterséges intelligencia és művészet az énképmérésben

Kellerwessel Klaus*  és Ujhelyi Adrienn

Eötvös Loránd Tudományegyetem Pszichológiai és Pedagógiai Kar, Pszichológiai Intézet, Budapest,
Magyarország

EREDETI KÖZLEMÉNY

Beérkezett: 2023. augusztus 21. – Elfogadva: 2024. március 18.

© 2024 A szerző(k)



Háttér és célkitűzések: A tetszőleges szövegi bemenet alapján művészi minőségű képeket alkotó mesterséges intelligencia (MI) megjelenése eddig még feltáratlan lehetőségeket tartalmazhat a pszichológia számára. Jelen kutatás során egy kvantitatív és kvalitatív módszereket szintetizáló és a résztvevő számára is élvezetes új énképmérő módszer kidolgozására tettünk kísérletet. A módszer elődjei közé tartoznak a különböző nyitott kérdéses önjellemzések, az önarcképfestés művészetterápiái tradíciója, illetve a számítógép-alapú testképvizsgáló eljárások. **Módszer:** A kutatás során arra kértünk 11 fiatal felnőtt résztvevőt, hogy a Midjourney nevű szoftver segítségével készítsenek képeket önmagukról. Az egyes képek alapjául szolgáló parancssorokat kategóriaelemzésnek vetettük alá, a résztvevőkkel pedig félig strukturált interjúkat vettünk fel. **Eredmények:** Az elemzések során szignifikáns egyéni különbségeket találtunk a résztvevők által említett önjellemzési kategóriákban, illetve a kategóriák időbeli változásában. Hogy ezek a különbségek milyen szituációs vagy személyen belüli hatásoknak tulajdoníthatók, eljövendő kutatások tárgyát fogja képezni, például a flow állapotra való egyéni érzékenység és az identitásállapot szerepének vizsgálatán keresztül. Az interjúk tanulsága szerint a résztvevők egy izgalmas önismereti tanulásként, identitásuk különböző aspektusaival való kísérletezésként élték meg a folyamatot. Többen megdöbbentek a gép képességeitől és saját művészi teremtőerejüktől, azonban az ember-gép kommunikáció félresiklásai több esetben megnehezítették az alkotást. **Következtetések:** Az általunk kidolgozott eljárás magában hordozza a potenciált arra, hogy a későbbiekben egy valid és reliábilis mérőeszközzé válhasson, nemcsak az énkép, hanem bármilyen attitűd mérésére. Az énhatékonyság-érzés megtapasztalása, illetve az önismereti tanulás érzete hosszú távon akár egy művészetterápiás felhasználásnak is az alapját képezheti.

* Levelező szerző. E-mail: kellerwesselklaus@gmail.com

KULCSSZAVAK

mesterséges intelligencia, művészet, énkép, önjellemzés, identitásállapot, módszertan

„Az én valószínűleg az általunk valaha létrejött
leglenyűgözőbb műalkotás, és minden bizonnyal
a legbonyolultabb”

Jerome Bruner

BEVEZETÉS

Nem vagyok robot

Edmond de Belamy portréja 432 500 dollárért kelt el egy londoni árverésen 2018. október 25-én. Ez a 147 396 000 forint¹ már csak azért is meglepő, mert a művész legelső elkészült képéről van szó. A kép persze nem önmagáért volt érdekes, hanem az alkotója miatt: ez volt az első mesterséges intelligencia (MI) által generált portré az emberiség történelmében (Du Sautoy, 2020).

Az első, önmagát MI művészettel foglalkozóként definiáló kiállítás az 1978-as londoni „Cybernetic Serendipity” volt. Ennek címében a „Serendipity” (ügyesség, szerencse) szó egy perzsa tündérmesére utal (Serendip három hercege, Goodman, 1961), ahol a főszereplők csodás véletleneknek hála folyton olyan dolgokra bukkantak, amelyeket nem is kerestek. Mintha a mesterséges intelligenciát csetlő-botló véletlenszám-generátornak láttuk volna, ami merő véletlenségből teremt néha értéket.

Az elmúlt 5 évben megszorodtak az önmagukat MI művészettel foglalkozóként hirdető kiállítások, címeiket pedig izgalmas összevetni a „Kibernetikus Véletlennel”. A 2014-es londoni „Creative Machine” (Kreatív gép) kiállítás (2014) névadója már humánspecifikusnak tartott kreativitást tulajdonít a gépeknek, a 2022-es budapesti: „Nem vagyok robot” (Ludwig Múzeum, 2022) kiállításcímében feltételezett megszólaló pedig már annyira antropomorf, hogy önmagát is antropomorfizálja. Mindezt a 2019-es londoni „AI: more than human” (MI: több, mint ember) címe (Wood, 2019) viszi el a logikai végkövetkeztetésig, azt sugallva, hogy még a humánspecifikusnak tartott tulajdonságokban is alulmaradunk.

Bár természetesen a kiállításcímek elsődleges célja, hogy kiugrasszák befogadjukat a komfortzónából, ezért szándékosan provokatívák és tudottan túlzóak, de pont ezért lehetnek indikátorai saját koruk félelmeinek és reményeinek (Sas, 2012). A 2010-es évek MI által létrehozott „művészete” a kiállításcímek sugallatai és néhány drágán elkelt portré ellenére még mindig inkább ügyes egzotikum volt csupán, azonban 2022-ben a sakk és a go játékok után (Du Sautoy, 2020) már a képzőművészetnek nevezett évezredes játékban is elvesztette címvédő szerepét az ember. A 2022-es Colorado State Fairt Jason Allen *Théâtre D’opéra Spatial* című alkotása nyerte meg idén nyáron. Még a zsűritagok számára is csak Allennek a díj odaítélése utáni Discord-posztjából derült ki, hogy a képet egy Midjourney nevű algoritmus készítette. Ez volt az első olyan alkalom, hogy MI alkotás hozzon el egy eredetileg emberi művészekre szabott díjat (Roose, 2022). „Nyertem, és nem szegtem meg semmilyen szabályt” – mondta Allen, kirobantva ezzel egy művészeti portálokon és versenyeken átívelő szabályozási hullámot (Harwell, 2022, 1).

¹2022. december végi árfolyamon.



Nemcsak a jogi szabályozást előzte meg az MI művészet gyors fejlődése, hanem a művészet-elméletet is (Liu és Tao, 2022). A témával foglalkozó kevés filozófus közül Tao (2022) az MI-n keresztül készített alkotásokat a tudatos emberi művészet és az öntudatlan természeti szépség között félúton helyezi el (hasonlóan egy kertész által tervezett és gondozott kerthez [Kellerwessel, 2023]). Tao szerint az MI által készített alkotások befogadásélményében még túl nagy szerepet játszik az „MI-által készített” címke, ami sokszor elvonja a figyelmünket a kép potenciális művészi értékeiről, azonban az esztétikai hozzáállásunk fejlődésével lehetségesnek tartja a művészet és technológia „új harmóniáját” (Tao, 2022, 1).

Bár, mint láttuk, az MI felhasználásával gyártott alkotások „igazi” művészet volta ma sem számít magától értetődőnek, és a technológia jogi és etikai vonatkozásai továbbra is kétségesek (Jiang és mtsai, 2023), az MI-művészetnek nevezett diszciplína egyre több helyen talál gyakorlati felhasználásra, például az építészetben (Ploennigs és Berger, 2023), oktatásban (Hemment és mtsai, 2023), üzleti életben (Mayahi és Vidrih, 2022), de akár még a régészetben is (Cobb, 2023).

Belamy portréja óta csupán öt év telt el, de a Midjourneyhez és DALL-E 2-höz hasonló szoftverek ma már néhány jellemző (prompt) megadása alapján képesek bármilyen kép létrehozására, tematikai és formátumbeli megkötés nélkül (Chatterjee, 2022). Ezek a programok minden megadott promptsorra négy különálló képet készítenek, amelyek más-más vizuális alapötlet mentén képzelik el a leírtakat (1. melléklet). A felhasználó ezután döntést hozhat az elkészült képekről, további variációkat vagy részletgazdagabb verziókat kérhet a neki tetsző példányokból, vagy a promptsor kisebb-nagyobb megváltoztatásával előről kezdheti a feladatot (lásd 2. melléklet), képgenerációról generációra tenyésztve ki így az elképzeléseihöz-izlésvilághoz illeszkedő alkotásokat.

Gépi tükrök

A pszichológia és informatika fejlődése a 20. század második felének folyamán szorosan összefonódott. A kognitív paradigma alapját az a hasonlat képezi, miszerint az emberi elme maga is egy információfeldolgozó eszköz, amely a világból érkező inputok alapján igyekszik kiszámítani a legadekvátabb viselkedés-outputot (Pléh és Lányi, 1984). A számítógép azonban nemcsak hasonlatként, hanem elemzési eszközként is meghatározza a pszichológia fejlődését (Garb, 2000). Első bevetése az MMPI/MMPI-2 kérdőívek automatizált kiértékelése volt (Fowler és Miller, 1969), ami a maga korában komoly ellenkezést váltott ki a papír-ceruza kiértékelés támogatói felől (Butcher és mtsai, 2004), azóta viszont természetesnek vesszük.

A számítógépek azonban nemcsak adattárolási és -kezelési segédeszközök. Az anorexia nervosa (AN) klasszikus értelmezése szerint az egyén saját testének torz észlelése vezet az önéheztetésbe torkolló fogyási kísérletekhez (American Psychiatric Association, 2013). A „Túl kövér vagyok” kijelentés azonban nemcsak tény-, de értékijelentés is (Cash és Deagle, 1997). Hogyan lehet szétszálalni az egyén testképében szerepet játszó érzékszervi-észlelési és affektív-kognitív kiértékelő faktorokat?

Leindler és munkatársai (2015) arra kérték a résztvevőket, hogy egy virtuális referenciatest arányait állítsák be olyanra, amilyenek a saját testüket jelenleg észlelik: önmaguk minél pontosabb tükrévé kellett alakítaniuk a számítógépet. A számítógép által behozott új testképmérisi lehetőségeknek hála, ma már tudjuk, hogy az anorexiával élők sajáttest-jellemzőinek becslési pontossága minőségileg nem különbözik az általános populáció sajáttest-jellemzőinek becslési pontosságától, a zavar kialakulásában pedig nem a saját test torz észlelése, hanem a vele



való elégedetlenség (P. L. Cornelissen és mtsai, 2013) és a kis testsúlykülönbségek iránti tanult érzékenység (K. K. Cornelissen és mtsai, 2015) játssza a legfőbb szerepet.

Messzire jutottunk, az MMPI első számítógépes kiértékelése óta: a számítógépek alig ötven év alatt furcsa kuriózumból a pszichológiai kutatások nélkülözhetetlen részévé váltak (Reynolds és mtsai, 2021). Belamy portréja óta csak öt év telt el, de a változás nem kevésbé forradalmi. Az MI-művészet áttöréseinek hála, képzeletünk vizuálisan leképezhetővé, legbelsőbb gondolataink pedig talán minden eddiginél kifejezhetőbbé váltak.

Jelen kutatás során arra kértük a résztvevőket, hogy a (MidJourney, 2023) szoftverrel készítsenek képeket önmagukról. A célunk azoknak az elemzési módszereknek a felderítése volt, amelyek képesek lehetnek a képkészítés individuális különbségeinek megragadására ebben az újfajta gépi tükörben.

Az új eljárás és a kapott eredmények részletes bemutatása előtt egy gyors betekintést szeretnénk nyújtani az önjellemzés-mérés néhány eddigi bevett módjába mind a kvantitatív, mind a kvalitatív tradíciókból, hogy lássuk, miképpen törekszik az általunk kidolgozott módszer ezek szintetizálására.

Egymás felé forduló arcok

A személyiségpszichológia az én tudománya, célja pedig kettős: az általános emberi belső működés mélyebb megismerése, illetve az emberek közötti különbségek, egyediségük és utánozhatatlanságuk megragadása. Ez a két cél párhuzamba állítható az én két jamesi aspektusával (Gale, 1999), a címkézővel, vagyis a megfigyelő alannal (ezt fejezi ki az angol *I* szó), illetve a címkézettel, vagyis a megfigyelés tárgyával (ezt fejezi ki az angol *Me* szó).

Ez a kétarcú én a pszichés folyamatok örök origója, ő közvetít a biológiai szükségletek, a társadalmi elvárások és a valóság nyújtotta lehetőségek között. A gyermekben először a megfigyelő és ítélező én fejlődik ki (Harter, 2003), akinek még nincs tudomása arról, hogy mások is megfigyelik és felcímkézik őt. Hogy ő nemcsak egy cselekvő alany, hanem mások számára ítéletek tárgya is, csak 6–7 éves korában kezd el tudatosodni benne, aztán néhány éven belül már képessé válik közvetlenül is megfigyelni önmagát. Az én két arca egymás felé fordul, megfigyelő és megfigyelt, címkéző és címkézett egybeesik, ezzel pedig megkezdődik a harc egy egységes identitásért (Bosma és Kunnen, 2001; Marcia, 1993). Az egyén önjellemzései innentől különösen nagy jelentőségre tesznek szert.

Az énkép az egyén önmagáról alkotott hiedelmeinek, meggyőződéseinek és érzelmeinek az összessége, mindaz, amit az egyén önmagával kapcsolatban igaznak tart (Baumeister, 1999). Az énkép vizsgálatának legközvetlenebb módját az önjellemzés-vizsgálatok jelentik.

McGuire és McGuire (1988) szerint minden önjellemzés-vizsgálat során felmerül egy trade-off helyzet: amikor a résztvevőnek előre megadott válaszlehetőségekkel gazdálkodva, fix dimenziók mentén kell elhelyeznie önmagát (pl. kérdőíves módszertan), akkor igazából a kutató emberképét mérjük, nem az egyén önmagáról alkotott képét. Ezzel szemben a nyílt kérdéses módszertannál a résztvevő maga választhatja ki önmaga megítélésének dimenzióit, ő szabja meg a játékszabályokat, és ez a szabályválasztás már önmagában sokat elárul róla. Ez persze időigényesebb és leterhelőbb a résztvevő számára, és nehezebben értelmezhető a kutatónak, viszont McGuire és McGuire (1988, 1) a reaktív módszerek érdemeinek elismerésével együtt is a nyílt kérdéses módszertan mellett foglalnak állást, mondván, hogy az általa biztosított információ-többlet egyszerűen megéri, visszakozni a tartalomelemzéstől pedig egyszerűen „gyávaság”.



A nyílt kérdéses módszertanok egyik legnagyobb kutatási hagyománnyal rendelkező példája az énfogalmak kulturális különbségeinek mérésére kidolgozott Húsz állítás teszt (Twenty Statements Test; (Koltói, 2016; Kuhn és McPartland, 1954). Ennek során a résztvevő feladata az, hogy húsz darab „Én ... vagyok” állításban foglalja össze mindazt, amit önmagáról gondol. Az állítások kódolása és kiértékelése az állítások kategóriákba sorolása mentén zajlik (lásd 3. melléklet).

Azt emeljük ki magunkról, ami minket is kiemel, McGuire és McGuire (1988) szerint az énkép megkülönböztethetőségi funkciója még a társas kívánatosságot is felülmúlja, ezért például az alacsony és túlsúlyos gyerekek is rendre megemlítik, hogy alacsonyak, illetve túlsúlyosak, amennyiben ez megkülönbözteti őket a többiektől.

Kívül és belül

Mind a McGuire és McGuire (1988) által reaktívnek (a résztvevő fix válaszlehetőségekre reagál), mind a nyílt kérdésesnek nevezett módszertanok eredendően kvantitatívak, így fő céljuk a számszerűsítés. Ez viszont problematikusá teszi a korábbi érvelést, mivel a résztvevő valószínűleg nem számokon keresztül fogja jellemezni önmagát. A Húsz állítás teszt során se maguk az elhangzott állítások válnak elemzés tárgyává, hanem a dimenziók, amelyek mentén elhangoztak. A résztvevő számára az a fontos, hogy ő kínai, nem pedig az, hogy a nemzetiségén keresztül jellemzi magát, az a fontos, hogy szerelmi bánata van, nem pedig az, hogy az affektív állapot volt az általa negyedikként felhozott kategória. A résztvevő a saját maga által választott önjellemzési dimenziók helyett a kutató által létrehozott metadimenziók mentén lesz kiértékelve.

A kvalitatív módszertanok adáz védelmezője (Norris, 2000) amellet érvel, hogy a pszichológiáról alkotott fogalmunk felszínes. Szerinte minden kutatás célja egy új szemléletmód elsajátítása, a világról szóló tudásunk értelmes egészé rendezése, ehhez pedig az adatok pusztán racionális megértése nem elég. Szerinte a kutatás során meg kell tanulnunk egyszerre látni kívülről és belülről is az adott problémát, és a prezentálás során meg kell próbálni a kutatás résztvevőinek belső élményvilágát is átadni.

Ugyan Norris eredetileg a saját drámaközpontú módszertana mellett érvelt, de az általa megszabott kritériumoknak eleget tehet az önarcképfestés évszázados művészetterápiái tradíciója is. A Gestalt-iskolához tartozó Perls már a 20. század elején használt önarcképfestésen alapuló művészetterápiás technikákat (Rhyne, 2001), az egyén önképének képként való leképezése pedig azóta bevett és elismert módszerré nőtte ki magát (Muri, 2007).

A képkészítés folyamatán, az elkészült képeken, és az utólagos megbeszélésükön keresztül nemcsak a terapeuta ismeri meg jobban az egyént, hanem ő is önmagát. Muri (2007) szerint a depresszióval küzdő egyének számára revelatív lehet az elkészült portrék visszanezése: tudomásul vehetik szomorúságukat anélkül, hogy eluralkodna rajtuk.

Muri (2007) azt is részletezi, hogyan nyilvánulhatnak meg különböző mentális zavarok vagy lelki események az elkészült portrékban: az Alzheimer-kór például hiányos test és megvágott fej ábrázolásában, a gyerekkori abúzus pedig kis, groteszk figurák szerepeltetésében, illetve fakó színek használatában.

Az önarcképek megbeszélése és értelmezése során a résztvevő által megszabott dimenziók (a kép tartalma, színvilága, az elkészítés folyamata, a benne kiváltott érzések) kerülnek előtérbe, így eleget tesznek a mcguire-i személyreszabottság-elvárásnak, az elkészült képeket pedig utólag ugyanúgy lehet prezentálni eredeti formájukban, mint egy színdarabot (lásd 10-14. melléklet), ez pedig norrisi tudományelsivárodás ellen hat.



Az énfogalom körvonalazódása

Az általunk kidolgozott eljárás során előállított adatmennyiség (képek, promptsorok, interjúk) több szinten, kvantitatív és kvalitatív módszertannal egyaránt elemezhető. Az MI művészeti szoftvereken keresztüli önképkeltés felfogható a klasszikus önarcképfestés (képek) és a Húsz állítás teszt (promptsorok) ötvözeteként.

Mindkét említett eljárás során egy résztvevő egy alkalommal egy outputot termel. Az MI művészszoftverek segítségével azonban egy résztvevő akár 40–50 képet is elkészíthet fél óra alatt, ráadásul minden egyes promptsor értelmezhető egy-egy önmagában teljes önjellemzőként, vagyis a képaradattal párhuzamosan potenciálisan 10–20 teljes önjellemző is elkészülhet.

Minden döntésnek (pl. egy már használt prompt bekerüljön-e a képek újabb generációjába?, tetszik-e az adott kép annyira, hogy variációkat kérjen le belőle?) verbális és vizuális lenyomata marad, vagyis a képek fokozatos átalakulásán keresztül talán magát az énfogalom körvonalazódását láthatjuk leképeződni.

Bár jelen módszer hosszú távon a két tradíció szintézisére törekszik, természetesen nem képes kiváltani egyiket sem. A Húsz állítás teszt előnyei az alacsonyabb idő- és erőforrásigény, illetve a bejáratott elemzési módok, az önarcképfestésnél pedig minden esetvonásban ott lesz a teljes ember (Edwards, 1987), ezért minden más módszernél mélyebb betekintést fog nyújtani az egyén szavakon túlmutató gondolataiba.

Jelen kutatás célja

Jelen kutatás célja annak a felderítése, hogy az egyén önmagáról alkotott hiedelmeinek rendszere (énképe) befolyásolja-e valamilyen formában a képgeneráló szoftver használata során hozott döntéseit és az ezáltal generált képi tartalmakat. Mivel még nem volt korábban kísérlet MI-vel készített alkotások pszichológiai mérésben való felhasználására, első lépésként csupán azt vizsgáljuk, hogy megjelennek-e egyének közötti különbségek a képkészítés mintázataiban, és hogy ezek a különbségek megragadhatóak-e az általunk javasolt kategóriaelemzési módszerrel.

Az ehhez kapcsolódó kutatási kérdések a következők:

1. Van-e résztvevők közötti különbség a felhasznált kategóriákban?
2. A felhasznált kategóriák idővel változnak-e?
3. Máshogy változnak-e idővel a felhasznált kategóriák résztvevőnként?

Az interjúk tartalomelemzése során a mesterséges intelligenciával való együttalkotás élményére, az egyénekre jellemző képkeltési stratégiákra és az elkészült képekről alkotott véleményekre vagyunk kíváncsiak.

MÓDSZER

Az eljárás menete

Az 50 perces alkalom során a vizsgálatvezető kettesben van a résztvevővel. A használt képgenerálási program (Midjourney[®]) működésének gyors ismertetése és az informált beleegyezés biztosítása után a vizsgálatvezető arra kéri a résztvevőt, hogy az elkövetkező fél óra során készítsen képeket magáról tetszőleges promplok megadásával, a sztenderdizált instrukció alapján (lásd 4. melléklet).



A résztvevők a képek elkészítése során technikai jellegű kérdésekkel bármikor fordulhattak a vizsgálatvezetőhöz, az esetleges nyelvi nehézségek enyhítésére pedig a DeepL nevű internetes fordítóprogramot használhatták. Más segítségük nem volt, nem használhattak például a telefonjukon található konkrét képet, sem pedig tükröt mintaként.

A félóra az első sikeresen elkészült képcsomag betöltése után indult, a letelte után pedig a résztvevő még befejezhette a már elkezdett promptsort, de újat már nem kezdhetett el.

A képkészítés fázisát egy félig strukturált interjú (Kovács és mtsai, 2019) követte (lásd 5. melléklet).

Etikai vonatkozások

A kutatásban való részvétel önkéntes volt, érte anyagi jutalmazás nem járt. A kutatásban 18 éven felüli, pszichiátriai kezelés alatt nem álló egyének vettek részt. A résztvevőknek lehetőségük volt bármikor, akár indoklás nélkül is megszakítani a részvételt. A kutatás folyamatáról, az adatvédelmi szabályozásról, illetve az elkészült képek és az interjú hanganyagának további sorsáról részletes tájékoztatást kaptak.

Jelen kutatást az ELTE PPK etikai bizottsága engedélyezte (az engedély száma: 2022/454). Az adatfelvétel 2022 novemberében zajlott.

A minta

A kutatásban részt vevő 11 kitöltő sok tekintetben homogén mintát képezett: 10 nő (90,91%) és 1 férfi (9,09%) volt közöttük, átlagéletkoruk: 22,64 év (SD = 0,98). Kilencük állandó lakhelye a főváros (81,82%), kettejüké pedig megyeszékhely (18,18%). Mindannyian egyetemisták. A 11 résztvevő összesen 154 képirási paranccsal nagyjából 650 képet termelt. A parancsok összesen 1588 promptból álltak, ebből 402 volt az új (az adott személynél először szereplő) információ.

KATEGÓRIAELEMZÉSEK

Módszertan

A statisztikai elemzésekhez a JASP 0.16.4 szoftvert használtuk. Az elemzés egysége az egyes prompt, a különböző változók pedig ehhez voltak hozzárendelve (lásd 6. melléklet). A következőkben először rátérünk a szövegfolyam promptokba szegmentálására, majd a felhasznált változókat, illetve a tervezett elemzéseket körvonalazzuk.

Promptokra bontás. Az elemzés során a felhasználó által írt természetes nyelvű bemeneteket manuálisan önálló jelentéssel bíró egységekre, vagyis promptokra bontottuk. Az egyszerű, szavakra való szegmentálással szemben azért volt szükség ezeknek az új jelentésegységeknek a kijelölésére, mivel a szavak jelentős része csak egy nagyobb egész részeként bír a jelen elemzés szempontjából is hasznos jelentéssel (például az olyan nyelvtani funkcióval rendelkező szavak, mint a jelen kutatásból származó „to” a „likes to party” promptban, vagy a „feels” a „feels himself free” kifejezésben). Más esetben a szavak egy kifejezés részeként teljesen új értelmet nyernek („big” a „Big Ben”-ben, vagy a „studying” a „studying psychology”-ban), és ez az általuk említett információ kategóriáját is érintheti (test vagy helyszín, illetve cselekvés vagy szerep – lásd a Kategóriák alfejezetben, illetve a 8. mellékletben).



A jelentésteli egységek megtalálásának problémájával (mint a legtöbb természetes nyelvfeldolgozó algoritmus) maga a Midjourney is szembesül, és a miénkhez valamelyest hasonló megoldást alkalmaz: a kiképzése folyamán megtanulja a betanítására szánt képekhez mellékelt leírásokat olyan egységekre bontani, amelyek mentén ezeket a képeket a legjobban elkülönítheti, később pedig a felhasználó által megadott parancsokat ezekkel a megtanult címkékkel közvetlenül összehasonlítható egységekre (token) tagolja (Xie, Pan, Ma, Jie és Mei, 2023).²

Az általunk alkalmazott manuális szegmentáció hátránya, hogy nem az egyes résztvevők, de nem is a Midjourney program által használt jelentésegység-határokat tükrözi vissza (mivel egyik sem rendelkezik fix határokkal, illetve mindkettő hozzáférhetetlen számunkra), előnye viszont, hogy rugalmas, sokrétű, és könnyen interpretálható elemzési lehetőségeket biztosít.

Kategóriák. Minden promptot egyetlen felsőbb kategóriába soroltuk a benne foglalt információ típusa szerint (7. melléklet). Az általunk javasolt teljes kategorizálási rendszer megtalálható a mellékletben (8. melléklet). A hierarchia csúcán álló két kategóriát a Gestalt-pszichológia nagy hagyománnyal rendelkező figura-háttér szétválasztása alapján hoztuk létre (Rhyné, 2001), a főbb alkategóriák pedig a következők (a jelen kutatásból származó példapromptokkal):

- **Figura**
 - Külső („very very curly hair”, „oily skin”, „no makeup”)
 - Belső („open-minded”, „likes to party”, „artist”)
 - Demográfia („23 years old”, „woman”, „European”)
- **Háttér**
 - Kép („painting”, „steampunk”, „dim lights”)
 - Helyzet („surrounded by mountains”, „dancing”, „feels himself free”)
 - Társaság („people around”, „sea creatures”)

Képsorszám. Minden egyes prompthoz csoporttagságként rendeltük hozzá a képek közötti elhelyezkedését, vagyis hogy hányadik képcsomaghhoz (promptsorhoz) tartozik.

Képötöd. Az egyes résztvevők által elkészített képeket egységesen öt részre bontottuk, így a kategóriaeloszlások időbeli változása a személyek által termelt eltérő mennyiségű képcsomagtól függetlenül is vizsgálhatóvá vált. Ezt az új változót nevezzük képötödnek.

JASP Képlete:

$$\text{Képötöd} = \text{cut} \left(\frac{\text{képsorszám}}{\max(\text{képsorszám}) | \text{résztvevő}}, 5 \right)$$

Az, hogy az elkészült képeket pont öt, és nem mondjuk három vagy tíz részre bontottuk, egy önkényes döntés volt, a mintaelemszámtól és a kívánt részletességtől függően állítható az egyes változók felbontása, akár képharmadra, akár képtizedre.

Új prompt. Ez egy igaz/hamis változó, amely azt vizsgálja, hogy az adott prompt először merült-e fel a résztvevő önjellemzésében. Az első képen minden prompt újnak számít. Ennek segítségével látható, hogy az adott információegység mikor merült fel először mint ötlet. Az összes prompttal való számolással szemben itt egyetlen információ sincs kétszer számolva.

²A természetes nyelvfeldolgozó algoritmusok által használt tokenek mérete modelltől modelltől változhat, elképzelhető, hogy szavaknál nagyobb és szavaknál kisebb tokenek is.



Az elemzések

Az elemzések célja annak körvonalazása lesz, hogy a különböző kategóriaemlétek milyen más vizsgált változók mentén mutatnak szignifikáns eltéréseket, ezért végig a nominális változók (kategóriák) más nominális/ordinális változók (résztevők, képötödök) mentén való eloszlását vizsgáló kontingenciátáblákat fogunk használni (lásd 10–12. melléklet). A khi-négyzet-próba azt vizsgálja, hogy az adott kontingenciátábla oszlopait képező változó befolyásolja-e szignifikánsan a táblázat sorait képező változó eloszlását (pl. befolyásolja-e a kategóriaemlétek eloszlását, hogy melyik képötödben járunk?).

1. *A kategóriaeloszlás személyközi különbségei.* Van-e különbség az egyes résztvevők által használt önjellemzési kategóriák között, ahogy ezt a Húsz állítás teszt mentén várhatjuk? A kérdés megválaszolásához használt kontingenciátábla sorait az egyes kategóriák, oszlopait pedig a résztvevők jelentik. Az elemzést különös jelentősége miatt a kategória-hierarchia több szintjén lefuttatjuk, képötödökre és promptötödökre bontva, összes promptokkal és csak új promptokkal külön is.
2. *A kategóriaeloszlás képötödközi változásai.* Van-e különbség a kategóriák előfordulásában az egyes képötödök között: tapasztalható-e a feladat során eltolódás az önjellemzés fókuszában? A kontingenciátábla sorait a kategóriák, oszlopait pedig a képötödök jelentik majd. A képötödökhöz kapcsolódó elemzéseket (2-es és 3-as) csak az új promptokon végezzük majd el, mivel azok vannak éppen aktuálisan a résztvevő fókuszában.
3. *A kategóriaeloszlások képötödközi változásainak személyközi különbségei.* Tapasztalhatók résztvevők közötti különbségek a képötödök közötti különbségekben? Ha feltételezzük, hogy ez egy önismereti folyamat, ahol mindenki más utat jár be, akkor elképzelhetőnek tűnik, hogy az önjellemzés fókusza máshogyan változzon az egyes egyéneknél. Hogy ez kvantitatívan is vizsgálhatóvá váljon, (a 9. mellékletben látható módon) először kigyűjtjük az egyes kategóriák adott képötödben belüli többi kategóriához viszonyított százalékos megoszlását az egyes résztvevőknél (ez a 2-es számú elemzés résztvevőkre bontott kontingenciátáblázatairól – 12. melléklet – válik majd leolvashatóvá), aztán kivonjuk az első képötöd százalékos megoszlásaiból az utolsó képötöd százalékos megoszlásait, megkapva ezzel a kategória eloszlásának kezdeti és végállapota közötti különbséget. A kategóriaváltozások személyek közötti különbségeit egy khi-négyzet-próbával hasonlítjuk össze, ahol a kontingenciátábla sorait az egyes kategóriák, oszlopait a résztvevők, az adott cellák esetszámát pedig az első és utolsó képötöd közti különbség jelenti. Mivel az elemzés szempontjából a különbség iránya nem, csak a mértéke fontos, ezért a különbség abszolút értékét vesszük.

Eredmények

1. *A kategóriaeloszlás személyközi különbségei.* Khi-négyzet-próba segítségével szignifikáns személyközi különbségeket találtunk a használt önjellemzési kategóriákban, a kategória-hierarchia mindhárom szintjén ($X^2 = 994,192$; $df = 50$; $p < 0,001$). Ez a különbség mind az összes prompt, mind az új promptok esetében megnyilvánult (1. táblázat). Az egyszerűség kedvéért innentől a 6 kategóriás felbontást fogjuk használni, mivel a figura-háttér elválasztásnál mélyebb betekintést nyújt, de nem válik a 20 kategóriához hasonlóan zavarossá. A 6 kategória menti személyközi eltérés majdnem minden képötödben szignifikáns (2. táblázat). A 6 kategória résztvevők közötti eloszláskülönbségeit a 10. és 11. melléklet szemlélteti.



1. táblázat. A kategóriaeloszlások személyközi különbségeit vizsgáló khi-négyszet-próbák három felbontásban, összes prompt és csak új promptok esetében is

	Összes prompt	Csak új prompt
<i>figura-háttér szint</i>	$X^2 = 357,513; df = 10; p < 0,001$	$X^2 = 357,513; df = 10; p < 0,001$
<i>6 kategóriaszint</i>	$X^2 = 994,192; df = 50; p < 0,001$	$X^2 = 249,862; df = 50; p < 0,001$
<i>20 kategóriaszint</i>	$X^2 = 1666,759; df = 150; p < 0,001$	$X^2 = 116,492; df = 150; p < 0,001$

2. táblázat. A kategóriaeloszlások személyközi különbségeit vizsgáló khi-négyszet-próbák mindegyik képötödben, összes prompt és csak új promptok esetében is

	Összes prompt	Csak új prompt
<i>1. képötöd</i>	$X^2 = 182,838; df = 50; p < 0,001$	$X^2 = 110,260; df = 50; p < 0,001$
<i>2. képötöd</i>	$X^2 = 208,881; df = 50; p < 0,001$	$X^2 = 85,321; df = 50; p < 0,001$
<i>3. képötöd</i>	$X^2 = 320,842; df = 50; p < 0,001$	Nincs adat
<i>4. képötöd</i>	$X^2 = 306,966; df = 50; p < 0,001$	Nincs adat
<i>5. képötöd</i>	$X^2 = 337,501; df = 50; p < 0,001$	$X^2 = 101,265; df = 50; p < 0,001$

- A kategóriaeloszlás képötödközi változásai.* A khi-négyszet-próba szignifikáns különbséget talált a képötödök között a használt önjellemzési kategóriákban ($X^2 = 62,250; df = 20; p < 0,001$; lásd [12. melléklet](#)).
- A kategóriaeloszlások képötödközi változásainak személyközi különbségei.* Amire már a képötödközi kategóriaváltozások során tapasztalt rendkívül nagy szórások is utaltak: a khi-négyszet-próba megerősítette ($X^2 = 1046; df = 50; p < 0,001$), hogy szignifikáns személyközi különbségek vannak a kategóriák változásában, vagyis az önjellemzés fókuszának eltolódásában. A [14. melléklet](#) négy véletlenszerűen kiválasztott résztvevőn keresztül nyújt betekintést ebbe a sokféleségbe.

INTERPRETATÍV FENOMENOLÓGIAI ELEMZÉS

Bár a promptokban felmerült információkategóriák kvantitatív vizsgálata fontos kiindulópontot jelenthet a további kutatások számára, azonban a Húsz állítás teszt bevett elemzéseihez hasonlóan szintén fix kiértékelési dimenziókat használ, vagyis a kutató emberképét méri, az egyén önmagáról alkotott képe helyett. Ezt most egy teljesebb kép érdekében az interjúk kvalitatív elemzésével egészítjük ki.

Módszer

Az általunk alkalmazott interpretatív fenomenológiai elemzési módszer célja (IPA) a vizsgált téma kapcsán felmerülő szubjektív élmények minél átfogóbb ábrázolása. A begépetelt interjúk elemzésének első lépése az interjúalanyok számára fontossággal bíró témák azonosítása, majd az egyes szövegrészeket csoportosítása a kapott témák mentén. Mindezt a kapott szöveg egységes narratívába fűzése követi, a kutató személyes értelmezésén keresztül (Alase, 2017).



Eredmények

Kreálok valamit, csak így szavakból! (Az élmény). A résztvevők összességében pozitív tapasztalatnak tartották a teszten való részvételt („kifejezetten nagyon érdekes volt”). Ennek forrása-ként többen a képkészítő programokban rejlő potenciál megtapasztalását emelték ki („Tök szuper, hogy vannak ilyen dolgok, és ennyi mindenre fel lehet ezeket használni”), illetve a saját művészetteremtő erejükre való rácsodálkozást („Kreálok valamit, csak így szavakból!”).

A képkészítési folyamatnak a kiszámíthatatlansága a felhasználókban izgalomérzetként jelent meg („Nagyon, nagyon tök jó, annyira izgalmas volt! Az nagyon tetszett, hogy mi lesz belőle a végén, mert nem tudtam elképzelni az elején, hogy a végén kijöjjön egy ilyen kép, ami [...] annyira hasonlít rám, és olyan, mint egy első benyomás rólam” (15. melléklet), „Jó volt látni, ahogy egyre jobban kitisztul.”).

Az önleképezés bár sokaknak kellemes élmény volt („Kicsit rossz napom volt, és jól esett ezt így átgondolni, hogy mi nyugtat meg, meg miket szoktam csinálni.”), sőt felszabadító („Élveztem, annyira, annyira élveztem! Zseniális feladat volt, [...] majdhogynem ilyen önmegvalósítás”), olyan is akadt, akinek inkább zavarbaejtő („Nem tudom, nyilván ez is része a felmérésnek, de tök nehéz így magamról ilyet csinálni, szóval hogyha mondjuk az lenne, hogy van egy fotó egy emberről, vagy egy ismerősről, akkor tök más dolgokat írnék be... [...] De lehet, hogy ez csak nekem fura. Még így is jó volt.”).

A géppel való kommunikáció nehézségeit többen kiemelték („Furcsa volt ezt így végigcsinálni, mert nehéz azt átadni így egy számítógépnek, hogy hogy képezed el magadat”). A helyzet egzotikussága miatt azonban a résztvevők számára ez az élmény is inkább érdekesnek vagy szórakoztatónak bizonyult („nyilván így szétszúrtak a dolgok. Ennek például három szeme lett... de ez közben meg nagyon vicces is volt”).

A fél óra kevésnek bizonyult, ezt szinte az összes résztvevő kiemelte („El tudtam volna még csinálni pár órán keresztül”), többen játékként utaltak közben a tesztfelvételre („Lehetne több az idő, mert ezzel így nagyon sok időt eltudnék játszani.”). A fél óra letelte után a legtöbben alig akarták abbagyni („Csak még egy kép!”).

Ezen sosem gondolkodtam így (Kettős tanulás). Azt már korábban kiemelttem, hogy az egyes promptoknak mint replikátoroknak pozitív megerősítést jelent, ha a résztvevő (a szelekciós ízlés) több képbe is beleteszi őket, de magának a résztvevőnek is pozitív megerősítést jelentenek az ízlésvilágát eltaláló képek, vagyis bizonyos értelemben kétirányú tanulási folyamatról van szó.

A résztvevő oldaláról ez egyrészt rátanulás a gép nyelvére („Ha megtalálnám azokat az igazán jó parancsszavakat [...] akkor az jönne ki, amit én akarok.”), másrészt pedig önismereti tanulás, az én szempontjából centrális vonások megtalálása („Furcsa volt így elgondolkozni, a szón amit mondtál, hogy önkép, hogy nem is tudom mik jelennek meg nekem így magamról, ezen sosem gondolkodtam így igazából. Mik jutnak magamról először eszembe? Mikor érzem jól magam? Mikor vagyok tényleg én, a legjobb állapotomban? Mi az amit a legjellemzőbbnek érzek a magam leírására?”). Ez a két tanulási folyamat nem vált el élesen egymástól, mivel a centrális vonások megtalálásával a géppel való kommunikáció is leegyszerűsödött („Kezdtünk egy ilyen Lisztferencesebb beütésű, tizennyolcadik században, akkor pedig nagyon gyorsan rájöttem két dologra. Egy: smiling. Kettő: on the beach, szóval akkor a tizennyolcadik századból nagyon gyorsan átugrottunk a beachboyra, és akkor így maradtam ebben a világban, és nagyon piciket változtattam csak. [...] szerintem ez végig maradt, ezt egy ilyen core-abb beállításomnak érzem. [...] Ezek a képek tényleg nagyon-nagyon tetszettek”).



A centrális vonások keresése és azok pontosítása több résztvevőben is tudatos stratégiaként működött („Először próbáltam a fontosabb dolgokat, mint a haj, [...], és azután a részleteket kidolgozni, például a bőröm, hogy milyen árnyalata van.”), néhányan megmaradtak egy-egy „alapmondat” mentén, és a fél óra folyamán végig azt igazgatták („ez az alapmondatom ez végig benne maradt, de így nem húztam ki igazán semmit belőle. Inkább csak toldottam-foldottam, leginkább csak variáltam.”). Ez a jellemzőségi sorrend nemcsak a képek között, de az egyes képek promptsorain belül is tudatosan jelen volt néhányaknál („Előre írtam azt, amit szerettem volna.”).

Bár többen fényképszerű élethűsége törekedtek („Az arcnál sok időt töltöttem azzal, hogy rám hasonlítson, megpróbáltam optimalizálni.”), de volt olyan résztvevő is, aki egy konkrét fényképet akart reprodukálni („Még benne voltam a képgenerálásban, amikor eszembe jutott, hogy amit én ebből az egészből ki akarok hozni, az egy real kép, az létezik, és azt akarom vizsgagenerálni. Csináltak rólam egy ilyen képet, vagy lehet, hogy ez csak az eltorzulása, lehet, hogy több képet csináltak rólam, és az emlékezetem egybegyúrta, de volt egy, amit vissza akartam kalibrálni.”).

Nem jött össze, hogy egy bálnával ússzak (Gépi intuíciók és sztereotípiák). A felhasználó által adott promptok természetesen nem fedhetik le a kép minden aspektusát. A hézagokat a Midjourney az adatbázisában tárolt több millió kép alapján próbálja meg kitölteni, például ha a felhasználó megnevez egy színt, akkor (további utasítás hiányában) a kép meg nem nevezett színei ehhez fognak illeszkedni („Ez egy jó tapasztalat volt [...], hogy [amikor azt írtam, hogy] piros vagy zöld [felsője legyen] akkor harmonizáló színeket csinált [a háttérben]”). Ez persze nemcsak a színeknél, hanem a kép bármely jellemzőjénél előfordulhat („Érdekes volt igazából az, hogy mi az, amit nem is kellett megemlítenem, például hogy barna hajam van. Nyilván az átlagnak barna haja van, gondolom valami statisztikák alapján csinálja ezeket.”).

Amikor ez a hézagkitöltés a résztvevők ízlésének megfelelően történt, az a Midjourney intuíciójaként is lecsapódhatott, hiszen „kitalálta magától”, amiről még a résztvevő se tudta, hogy írni akar („Ebben így látom magam, és nagyon szép. Úgy érzem, hogy nem adtam meg sok támpontot, és nagyon jól megéreztem valamit, amit én elképzeltem. Ez annyira menő.” [16. melléklet]), „Először még gondolkodtam, hogy az arcszörzettel játszak egy kicsit, de aztán magától belejött ez a kis borosta, amit nem írtam bele, de belegondoltam, hogy egy kis borosta jól jönne, és tőkre megörültem, hogy halványan megjelent”).

Viszont amikor nem feleltek meg a szelekciós ízlésnek, akkor „sztereotípiaként” is lecsapódhatnak („Olyan, minthogyha lennének neki sztereotípiái: beírtam, hogy kedves és segítőkész és rögtön ilyen babaarcú lett”). Sokan túlidealizáltak, „túl szép”-nek találták ezeket a figurákat, főleg a tanulási folyamat elején („Amikor azt írom, hogy mogorva, akkor is ilyen... szép lány, és cukin mogorva...”, „Néztem, hogy én azért ennyire nem vagyok macsó és jóképű. Ez egy wannabe Alex.” [17. melléklet]), sokan még az inspirációt is látni vélik („ilyen kedves-pirososzás, nagyon ilyen Disney-hercegnős”, „Anime karakterek”). Más esetekben viszont az elkészült képek a groteszk felé torzultak („úristen, remélem nem így nézek ki” [18. melléklet]).

Többen kitartó harcot vívtak a kép egy-egy jellemzőjének megváltoztatásáért, de a gép ebben túl makacsnak bizonyult („Annyira ez van, gondolom, az ő tudásában hogy a vörös [haj] az mindenképpen természetes vörös, és hiába írtam bele, hogy festett vörös, az narancssárga maradt, és az volt bennem, hogy biztos be lehet ezt úgy valahogy írni, hogy azt ő is értse, de nem igazán sikerült. Ez volt a legnagyobb hiba, de amúgy egészen jól lefestett.”), „Itt viszont konkrétan leírtam, hogy egy bálnával szeretnék úszni, és nem jött össze, hogy egy bálnával ússzak.”). Ezt néhányan a



gép képzeletén belüli „összeférhetetlenségnek” tulajdonították („Nem vette [figyelembe] azt, hogy people in the background, itt, a könyvtárban. Szerintem az összeférhetetlen neki, hogy emberek legyenek a könyvtárban.”).

Néhány olyan jelenség azonban, ami elsőre gépi macacsságnak tűnt, utólag már a résztvevő saját figyelmetlenségeként kereteződött: előfordultak beleértett utasítások („Szerintem mi nem ugyanarra a könyvtárra gondoltunk, én inkább ilyen klubhelyésre gondoltam a Harry Potterben, de ezt nyilván nem tudta így kitalálni”), vagy odavetett paraméterek („beírtam azt, hogy zöld, amit nagyon szeretek, ami után már így zöldre csinálta a hajamat, ez pedig szerintem egy fika itt az orrában”).

Nemcsak a túl keveset változó képjellemzők, de a túl hirtelen változók is kihívást jelentettek („Láttam, hogyha beírom, hogy szeplő, akkor ilyen nagyon szeplősöket hoz ki, ha meg beírom, hogy bőrhiba, akkor meg nagyon sok lesz a bőrhiba.”).

Belső boldogság (A kép mint énképezés). A résztvevők igyekeztek személyes történeteikkel megtölteni a képeket, például a háttérben („ez [egészen olyan lett], mint a szobám”), a kellékeken („open wide shirt, long necklace, de hogy ezek létező dolgok, szóval itt van a nyakamon ez a necklace és hasonlók.”), vagy a képen jelen lévő egyéb szereplőkön keresztül („Szerintem ezek [a képek] is iszonyatosan jók lettek. Egy kicsit be akartam hozni ezeket a közösségi dolgokat is. Azt szerettem volna, hogy az látszódjon, hogy nők tűz körül táncolnak, nem tudom miért ez jutott eszembe, azt szerettem volna, hogy másokat is behozzak ebbe az önjellemzésbe, és így ez jutott eszembe” [19. melléklet]).

Az egyik résztvevő életének több korszakát is felfedezte a képeken („Ez is lehettem volna én 18 évesen, amikor hátrakötöttem a hajamat és egy tincs mindig maradt valahol elől vagy oldalt, valahol mindig kikószált.”, „megtaláltam egy teljesen olyan korszakomat, ami régebben volt, amikor a barátnőm levágta a hajamat, és akkor rövid lett. És volt egy ilyen beaches időszak is, amikor tényleg így néztem ki a beacheen, még Portugáliában.”), sőt, megtalálta még a jövőbeli énjét is („Nagyon vicces, mert hogy ezek énidetitásbelileg így teljesen [stimmelnek], de hogy ez a macsobb inkább [az életem] egy későbbi szakaszába illik.”). Az interjú során egy teljes éntablót állított fel az elkészült képekből („Életutat beívelve [sorbarendezhető]).

A képek jól sikerültnek értékelésében az általános hangulat játszott a legfőbb szerepet („Ez volt az áttörés, ez nagyon tetszett ez a kép, ennek a hangulatával nagyon tudok azonosulni. Ez nem is egy ilyen ideális én, hanem hogyha most átraknám magam egy ilyen virtuális valóságba, akkor valószínűleg így öltöztetném fel magam, vagy valami ilyesmi”, „Nekem ez a melankolikus hatás nagyon komfortos, szóval nem kell rám hívni a pszichológusokat”, „Volt egy-kettő, amire így külsőleg egyáltalán nem érzem, hogy hasonlítana, de hogy így belsőleg, a vibe-ja ennek teljesen. Itt variációkat is kértem rá egyből. Ezt is teljesen átéreztem, ez egy ilyen belső boldogság.”).

DISZKUZZIÓ

Az eredmények összefoglalása

Ahogy Belamy portréja, úgy ez a kutatás sem önmagáért fontos, hanem mint az eljövendő lehetőségek előfutára. A kategóriaelemzések tanulsága szerint a résztvevők által használt önjellemzési kategóriák, és a kategóriák időbeli változásai is szignifikáns személyközi különbséget mutatnak. Az, hogy ezek a különbségek milyen szituációs vagy személyen belüli hatásoknak tulajdoníthatók, eljövendő kutatások tárgyát kell, hogy képezze. Néhány ötletet már a diszkuszióban felvetünk.



Önmagunk megtalálása

Több résztvevő kiemelte, hogy tudatosan az énje centrális vonásainak megtalálására törekedett a feladat során („*Mikor vagyok tényleg én?*”). A 3-as számú elemzés arra utal, hogy szignifikáns különbség található az egyes résztvevők által bejárt fókuszutak között. [Marcia \(1993\)](#) identitás-állapot-modellje szerint a valódi identitás megtalálásához a serdülőnek végig kell mennie a potenciális szerepekkel való kísérletezések (exploráció), illetve a megfelelőnek ítélt szerepek melletti elköteleződés folyamatán, ha pedig mindez megtörténik, akkor a lehetőségek káoszából kiválik egy nagyjából konzisztens én. Egyes elméletalkotók ([Bosma és Kunnen, 2001](#)) az alábbi folyamatot nem tartják a serdülőkorra specifikusnak, hanem a felnőtt életút folyamán újra meg újra fellazuló és megszilárduló identitásról, az exploráció és elköteleződés hullámzó ciklusairól beszélnek. Vajon mennyiben bizonyul analógnak a Midjourney gépi tükrében történő önképezés folyamata az önmagunkban történő önképezés, vagyis az identitás megtalálásának a folyamatával? Az első pillanatok kaotikus zavarodottsága a promptokkal (amelyek az énről szóló információk, öncímkék) való kísérletezések (exploráció) során „*egyre jobban kitisztul*”, és ideális esetben a reprezentatívnak érzett promptok és képek melletti elköteleződéssel zárul („*Ezt [...] teljesen átéreztem*”).

Megkülönböztethető-e egymástól a kategóriák egyéni változásai mentén az elköteleződés nélküli exploráció örök bolyongása (moratórium), a túl rövidre zárt exploráció (korai zárás), illetve az elköteleződést és az explorációt is nélkülöző kapaszkodók nélküli diffúz identitásállapot ([Marcia, 1993](#))? Ezt a kutatás következő hullámában a [Balistreri és munkatársai, 1995](#) által kidolgozott Identitás-folyamat kérdőívvel fogjuk majd mérni.

Az alkotás sodrása

Egy pszichológiai teszt élvezhetősége fontos szempont a fejlesztés folyamán, nem pusztán etikai ([Koocher és Keith-Spiegel, 2008](#)), de a teszt validitásának szempontjából is ([Lumsden és mtsai, 2016](#)). Az interjúk tanulsága szerint a résztvevők élvezetesnek találták az eljárást. A művészi teremtőképességükre való rácsodálkozás az énhatékonyság-érzés megnövekedéséhez vezethetett, hasonlóan, mint az önarcképfestés során ([Muri, 2007](#)).

A bevonódást elősegítette még a folyamat „*izgalmas*” kiszámíthatatlansága is. A résztvevő a diffúziós folyamat végéig nem tudhatta, hogy a beírt promptsor eredményei az önképének megfelelő, pozitív megerősítésként ható képek lesznek-e, ez pedig egy változó arányú megerősítési tervként, lebilincselő erővel hathatott („*nem tudtam elképzelni az elején, hogy a végén kijöjjön egy ilyen kép, ami [...] ennyire hasonlít rám*”). Az önképekkel való kísérletezésnek ez a fajta optimális kiszámíthatatlansága kihívássá tette a résztvevők számára a feladatot, aminek köszönhetően többekben megvalósulhatott a flow-állapot alapjául szolgáló kihívás-készség egyensúly is ([Csikszentmihályi, 1997](#)). Az állapothoz kapcsolódó időélmény-torzulásra utalhatnak az olyan kifejezések is, mint: „*El tudtam volna még csinálni pár órán keresztül*”.

Az átélt flow-állapot mérésére a kutatás következő fázisaiban a [Magyaródi és munkatársai \(2013\)](#) által fejlesztett Flow-állapot kérdőívvel, az egyéni flow-átélési és élményben elmerülési készség mérésére pedig a Tellegen-féle adszorpció skálával ([Simor és mtsai, 2011](#)) kerül majd sor.

Az áramlatélmény és az énhatékonyság-érzés esetleges megtapasztalása, illetve a „Tanultam valamit magamról” érzete hosszú távon akár egy művészetterápiás felhasználásnak is az alapját képezheti valamilyen formában.



Potenciális gyenge pontok

Ahogy nem léphetünk kétszer ugyanabba a folyóba, úgy nem kaphatjuk meg kétszer pontosan ugyanazokat a képeket, hiába írjuk be ugyanazt a promptsort, és ez komoly reliabilitásbeli problémákat okozhat. Itt a résztvevő szelekciós ízlése jelentheti a kiutat, az, hogy döntést hoz mind az 50 elkészült képről és mind a 200 felhasznált promptról, és az őt nem jól leképező információk kiszekeltálódnak egy jól dokumentált folyamat során.

A módszer validitására pedig többek között a gépbe épített sztereotípiák jelentenek veszélyt. Az intelligenciák rugalmatlansága, hogy az eddigi ismereteiknek ellentmondó információt kezeljék, nemcsak az emberi társadalmakban, de több MI alapú technológia esetében is komoly problémákat okoz (Vespignani, 2020). A Midjourney fejlesztőinek szerencsére fókuszában van az efféle rugalmatlanság. 2022 áprilisában fontos mérföldkőnek számított a lovagló asztronauta képe (Heaven, 2022), amelyet korábbi programok nem tudtak volna elkészíteni, mivel a beléjük táplált adatokban nem találhattak lovagló asztronautákról szóló mintát, vagyis az a sztereotípiát élte bennük az asztronautákról, hogy nem tudnak lovagolni. Még további áttörésekre lesz szükség, hiszen (a festett vörös hajon kívül) egy asztronaután lovagló ló elkészítése még ma is komoly nehézségekbe ütközik (Marcus, 2022). Amikor először vázoltuk fel a kutatási ötletünket, még az a veszély fenyegette, hogy az életszerűtlen arcábrázolás fogja ellehetetleníteni, de egy héten belül ki is jött egy új fejlesztés, amely már ijesztően élethűvé tette a Midjourney arcait (Monge, 2022). Bízunk benne, hogy ahogy a jelen kutatás, úgy a Midjourney alapjául szolgáló technológia is hamarosan ki fogja nőni gyermekbetegségeit.

Amennyiben mégsem leszünk tanúi sem hirtelen áttöréseknek, sem lassú mérséklődésnek, a program sztereotípiáinak feltérképezése még mindig elégséges megoldás lehet. Ha ismerjük a programot, és látjuk, hogyan tud együtt dolgozni vele a résztvevő, akkor jobban megismerjük magát a résztvevőt is. Tudatos reflexióval ellensúlyozhatjuk a beépített torzításokat, pont mint az emberi intelligenciák esetében.

További kutatási irányok

Ugyan, mint láttuk, nagy lehetőségek rejlenek pusztán az önleképezések folyamatának vizsgálatában is, azonban az itt kidolgozásra kerülő eljárás tágabb keretek között is alkalmazható lehet. Az önmagunkról szóló attitűd bizonyos tekintetben ugyanolyan, mint a többi. Arra is kérhetjük a résztvevőt, hogy a Midjourney segítségével másokról alkosson képeket (mint erre az egyik interjúalany már utalt is). Dinamikus szempontból érdekes lehet például az egyén anyjáról/apjáról alkotott képeinek tanulmányozása, vagy mondjuk a párok egymásról alkotott képeinek összevetése. Természetesen a már létező attitűdtárgyak leképezéséhez sem kell ragaszkodni, téma lehet az egyén saját maga ideális, és mások által elvárt énjének összehasonlítása, vagy egy vágyott partner megalkotása is. Nem kell ragaszkodni a személyleképezéshez sem, a marketingkutatás (Sas, 2012) a különböző termékek vagy márkaimázsok tanulmányozására is használhatja a módszert.

LIMITÁCIÓK

Ahogy maga az énfogalmunk, vagy az MI művészet, úgy az énfogalmunkat vizsgáló MI-művészet sem teljes fegyverzetében pattant elő. Az alábbiakban bemutatott eljárás még az exploráció fázisában van, alig ébredt tudatára saját lehetőségeinek. Mibenlétének lassú körvonalazása



kísérletező-szellemű kutatások hosszú sorát igényli majd, de hisszük, hogy hamarosan egy valid és reliábilis eszközt tarthatunk majd a kezünkben.

Ezeknek az eljövendő kutatásoknak túl kell lendülniük a jelen kutatás hiányosságain: a résztvevők nemi, életkori és kulturális homogenitásán, a promptok szegmentálásának manuális voltán, illetve a külső validitás mérésének elhanyagolásán. Kreatív megoldásokat kell találniuk az eljárást alapjaiban fenyegető problémák megszelídítésére, mint például a képzés szten-derdizálhatatlan volta, illetve a gép belső sztereotípiái esetében.

ÖSSZEFOGLALÁS

A 2010-es évek elején piacra kerülő grafikai szoftverek lehetővé tették, hogy egy fokkal jobban megismerjük önünk egyik fontos szegletét, a testképünket. Jelen kutatás eredményei arra utalnak, hogy a grafikai szoftverekben történt újabb forradalom utat nyithat az énkép átfogóbb vizsgálati számára is.

Bruner az ember által létrejött legbonyolultabb és leglenyűgözőbb műalkotásnak hívta az ént (Du Sautoy, 2020), jelen módszer pedig ezzel összhangban műalkotások sorozatán keresztül igyekszik megragadni őt, és a folyamatot, ahogy megragadja önmagát.

Ez a sokszínű gépi tükör túllép a művészetten keresztüli önképezés eredendően kvalitatív voltán: a kategóriaelemzés általunk javasolt módszertana már jelenlegi formájában is kellően érzékenynek bizonyult (vagyis képes volt szignifikáns különbségek megragadására) több dimenzió mentén. Az, hogy pontosan milyen tényezőknek tulajdoníthatók ezek a résztvevők és kép-ötödök közötti különbségek, vagyis milyen pszichológiai konstruktumok megragadására válhatnak alkalmassá, további kutatások tárgya lesz majd.

A technológia fejlődésével valószínűleg egyre nagyobb teret kapnak majd életünkben a stabil diffúziós modellek, és a verbális parancsokon, promptokon keresztüli teremtés. Képzletünk formába öntése egyszerű kuriózumból talán pont úgy válik majd mindennapjaink nélkülözhetetlen részévé, mint az elmúlt 50 év alatt a számítógépek. Ha mi, pszichológusok már most elkezdünk a mélyére ásni annak, mi is zajlik le fejünkben a megfelelő kulcsszavak kiválasztásakor, ha már most elkezdünk modelleket építeni arra, hogyan tanul egymáshoz ember és gép, organikus és mesterséges intelligencia, akkor talán felkészültebben érhetik majd az emberi társadalmakat a jelen technológiával jött változások. Talán még a Tao által felvetett „új harmóniához” (2022, 1) is közelebb kerülünk.

A cél tehát most is, mint annyiszor, hogy az informatika tudományának forradalma ismét átterjedhessen a pszichológiára is (Kuhn és mtsai, 1984).

MELLÉKLET

A cikkhez tartozó melléklet elérhető online: <https://doi.org/10.1556/0016.2024.00077>.

IRODALOM

Alase, A. (2017). The interpretative phenomenological analysis (IPA): A guide to a good qualitative research approach. *International Journal of Education and Literacy Studies*, 5(2), 9. <https://doi.org/10.7575/aiac.ijels.v5n2p.9>.



- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Association.
- Balistreri, E., Busch-Rossnagel, N. A., & Geisinger, K. F. (1995). Development and preliminary validation of the ego identity process questionnaire. *Journal of Adolescence*, 18(2), 179–192. <https://doi.org/10.1006/jado.1995.1012>.
- Baumeister, R. F. (1999). *The self in social psychology*. Psychology Press.
- Bosma, H. A., & Kunnen, E. S. (2001). Determinants and mechanisms in ego identity development: A review and synthesis. *Developmental Review*, 21(1), 39–66. <https://doi.org/10.1006/drev.2000.0514>.
- Butcher, J. N., Perry, J., & Hahn, J. (2004). Computers in clinical assessment: Historical developments, present status, and future challenges. *Journal of Clinical Psychology*, 60(3), 331–345. <https://doi.org/10.1002/jclp.10267>.
- Cash, T. F., & Deagle, E. A. (1997). The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: A meta-analysis. *International Journal of Eating Disorders*, 22(2), 107–126. [https://doi.org/10.1002/\(SICI\)1098-108X\(199709\)22:2<107::AID-EAT1>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1098-108X(199709)22:2<107::AID-EAT1>3.0.CO;2-J).
- Chatterjee, A. (2022). Art in an age of artificial intelligence. *Frontiers in Psychology*, 13, 1024449. <https://doi.org/10.3389/fpsyg.2022.1024449>.
- Cobb, P. J. (2023). Large language models and generative AI, Oh My!: Archaeology in the time of ChatGPT, Midjourney, and beyond. *Advances in Archaeological Practice*, 11(3), 363–369. <https://doi.org/10.1017/aap.2023.20>.
- Cornelissen, K. K., Bester, A., Cairns, P., Tovée, M. J., & Cornelissen, P. L. (2015). The influence of personal BMI on body size estimations and sensitivity to body size change in anorexia spectrum disorders. *Body Image*, 13, 75–85. <https://doi.org/10.1016/j.bodyim.2015.01.001>.
- Cornelissen, P. L., Johns, A., & Tovée, M. J. (2013). Body size over-estimation in women with anorexia nervosa is not qualitatively different from female controls. *Body Image*, 10(1), 103–111. <https://doi.org/10.1016/j.bodyim.2012.09.003>.
- Creative Machine (2014). Letöltve 2024. 03. 11-én: <http://creativemachine.org.uk/>.
- Csikszentmihályi, M. (1997). *Flow – Az áramlat. A tökéletes élmény pszichológiája*. Budapest: Akadémiai Kiadó.
- Du Sautoy, M. (2020). *The creativity code: Art and innovation in the age of AI*. Harvard University Press. <https://doi.org/10.4159/9780674240407>.
- Edwards, B. (1987). *Drawing on the artist within: A guide to innovation, invention, imagination, and creativity*. Simon and Schuster.
- Fowler, R. D., & Miller, M. L. (1969). Computer interpretation: Its use in clinical practice. *Archives of General Psychiatry*, 21(4), 502. <https://doi.org/10.1001/archpsyc.1969.01740220118014>.
- Gale, R. M. (1999). *The divided self of William James*. Cambridge University Press.
- Garb, H. N. (2000). Computers will become increasingly important for psychological assessment: Not that there's anything wrong with that. *Psychological Assessment*, 12(1), 31–39. <https://doi.org/10.1037/1040-3590.12.1.31>.
- Goodman, L. A. (1961). Notes on the etymology of serendipity and some related philological observations. *Modern Language Notes*, 76(5), 454. <https://doi.org/10.2307/3040685>.
- Harter, S. (2003). The development of self-representations during childhood and adolescence. In M. R. Leary, & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 610–642). The Guilford Press.
- Harwell, D. (2022). He used AI to win a fine-arts competition. Did he cheat? *The Washington Post*. Letöltve 2024. 03. 11-én: <https://www.washingtonpost.com/technology/2022/09/02/midjourney-artificial-intelligence-state-fair-colorado/>.



- Heaven, W. D. (2022). This horse-riding astronaut is a milestone on AI's long road towards understanding. *MIT Technology Review*. Letöltve 2024. 03. 11-én: <https://www.technologyreview.com/2022/04/06/1049061/dalle-openai-gpt3-ai-agi-multimodal-image-generation/>.
- Hemment, D., Currie, M., Bennett, S., Elwes, J., Ridler, A., Sindors, C., ... Warner, H. (2023). AI in the public eye: Investigating public AI literacy through AI. In *FACCT'23: Proceedings of the 2023 ACM conference on fairness, accountability, and transparency* (pp. 931–942). New York: Association for Computing Machinery (ACM). The 2023 ACM Conference on Fairness, Accountability, and Transparency, Chicago, Illinois. <https://doi.org/10.1145/3593013.3594052>.
- Jiang, H. H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., ... Gebru, T. (2023). AI art and its impact on artists. In *Proceedings of the 2023 AAAI/ACM conference on AI, ethics, and society* (pp. 363–374). <https://doi.org/10.1145/3600211.3604681>.
- Kellerwessel, K. (2023). Angyal a gépben. A mesterséges intelligencia veszélyei és lehetőségei az irodalomban. *Litera*. Letöltve 2024. 03. 11-én: <https://litera.hu/irodalom/elso-kozles/angyal-a-gepben-a-mesterseges-intelligencia-veszelyei-es-lehetosegei-az-irodalomban.html>.
- Koltói, L. (2016). „Ki vagyok én?” – A pedagógusidentitás mérése a Húsz állítás magamról c. teszttel = „Who am I?” Using the twenty statement test to measure professional identity of student teachers. *Gradus*, 3(1), 142–147.
- Koocher, G. P., & Keith-Spiegel, P. (2008). *Ethics in psychology and the mental health professions: Standards and cases*. Oxford University Press.
- Kovács, A., Kiss, D., Kassai, S., Pados, E., Kaló, Z., & Rácz, J. (2019). Mapping qualitative research in psychology across five Central-Eastern European countries: Contemporary trends: A paradigm analysis. *Qualitative Research in Psychology*, 16(3), 354–374. <https://doi.org/10.1080/14780887.2019.1605271>.
- Kuhn, T. S. (1984). *A tudományos forradalmak szerkezete*. (Ford. Bíró Dániel). Budapest: Gondolat.
- Kuhn, M. H., & McPartland, T. S. (1954). *Twenty statements test* [Database record]. APA Psyc Tests. <https://doi.org/10.1037/t05100-000>.
- Leindler, M., Benkő, Z., Lapis, F., Léder, Á., Molnár, K., & Szalai, D. (2015). Egy háromdimenziós modellező és grafikai program, a DAZ Studio alkalmazásának lehetőségei a testkép vizsgálatában. *Mentálhigiénié és Pszichoszomatika*, 16(3), 253–270. <https://doi.org/10.1556/0406.16.2015.3.3>.
- Liu, W., & Tao, F. (2022). Art definition and accelerated experience: Temporal dimension of AI artworks. *Philosophies*, 7(6), 127. <https://doi.org/10.3390/philosophies7060127>.
- Ludwig Múzeum (2022). Nem vagyok robot. Letöltve 2024. 03. 11-én: https://www.ludwigmuseum.hu/system/files/news/attachments/2022-11/221103_nemvagyokrobot_kiadvany_100x180_online.pdf.
- Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games*, 4(2), e11. <https://doi.org/10.2196/games.5888>.
- Magyaródi, T., Nagy, H., Soltész, P., Mózes, T., & Oláh, A. (2013). Egy újonnan kidolgozott Flow Állapot Kérdőív kimunkálásának és pszichometriai jellemzőinek bemutatása. *Pszichológia*, 33(1), 15–36. <https://doi.org/10.1556/Pszicho.33.2013.1.2>.
- Marcia, J. E. (1993). The ego identity status approach to ego identity. In *Ego identity: A handbook for psychosocial research* (pp. 3–21). New York: Springer. https://doi.org/10.1007/978-1-4613-8330-7_1.
- Marcus, G. (2022). Horse rides astronaut. *Substack*. Letöltve 2024. 03. 11-én: <https://garymarcus.substack.com/p/horse-rides-astronaut>.
- Mayahi, S., & Vidrih, M. (2022). *The impact of generative AI on the future of visual content marketing*. ArXiv, abs/2211.12660. <https://doi.org/10.48550/ARXIV.2211.12660>.



- McGuire, W. J., & McGuire, C. V. (1988). Content and process in the experience of self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. XXI, pp. 97–144). Elsevier. [https://doi.org/10.1016/S0065-2601\(08\)60225-7](https://doi.org/10.1016/S0065-2601(08)60225-7).
- Midjourney (2023). *Midjourney (V5) [Text-to-image model]*. [Software]. <https://www.midjourney.com/>.
- Monge, J. C. (2022). MidJourney's major update fixes its biggest flaw. *Medium*. Letöltve 2024. 03. 11-én: <https://medium.com/geekculture/midjourneys-major-update-fixes-its-biggest-flaw-86dd0b353126>.
- Muri, S. A. (2007). Beyond the face: Art therapy and self-portraiture. *The Arts in Psychotherapy*, 34(4), 331–339. <https://doi.org/10.1016/j.aip.2007.05.002>.
- Norris, J. (2000). Drama as research: Realizing the potential of drama in education as a research methodology. *Youth Theatre Journal*, 14(1), 40–51. <https://doi.org/10.1080/08929092.2000.10012516>.
- Pléh, Cs., & Lányi, G. (1984). A „kognitív forradalom” és a magyar pszichológia. *Valóság*, 7, 13–28.
- Ploennigs, J., & Berger, M. (2023). AI art in architecture. *AI in Civil Engineering*, 2(1), 8. <https://doi.org/10.1007/s43503-023-00018-y>.
- Reynolds, C. R., Altmann, R. A., & Allen, D. N. (2021). *Mastering modern psychological testing: Theory and methods*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-59455-8>.
- Rhyne, J. (2001). The gestalt approach to experience, art, and art therapy. *American Journal of Art Therapy*, 40(1), 109.
- Roose, K. (2022). An A.I.-Generated picture won an art prize. Artists aren't happy. *The New York Times*. Letöltve 2024. 03. 11-én: <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>.
- Sas, I. (2012). *Reklám és pszichológia a webkorszakban*. Budapest: Kommunikációs Akadémia Könyvtár.
- Simor, P., Köteles, F., & Bódizs, R. (2011). Submersion in the experience: The examination of the Tellegen Absorption Scale in an undergraduate university sample. *Mentálhigiéné és Pszichoszomatika*, 12(2), 101–123. <https://doi.org/10.1556/Mental.12.2011.2.1>.
- Tao, F. (2022). A new harmonisation of art and technology: Philosophic interpretations of artificial intelligence art. *Critical Arts*, 36(1–2), 110–125. <https://doi.org/10.1080/02560046.2022.2112725>.
- Vespignani, A. (2020). *A jóslás algoritmusa: Hogyan befolyásolható a jövő a tudomány segítségével*. Libri.
- Wood, C. (2019). *AI – More than human*. London: Barbican Art Gallery.
- Xie, Y., Pan, Z., Ma, J., Jie, L., & Mei, Q. (2023, April). A prompt log analysis of text-to-image generation systems. In *Proceedings of the ACM Web Conference 2023* (pp. 3892–3902).

„I create something, from words alone!” – Artificial intelligence and art in self-image assessment

Klaus Kellerwessel and Adrienn Ujhelyi

Background and Objectives: The new artificial intelligence (AI) based methods that enable users to create high quality images based on natural language textual inputs alone could pose significant potential for psychology. This research aimed to develop a novel self-image assessment method by synthesizing both quantitative and qualitative approaches and ensuring an engaging experience for participants. Antecedents of this method include various open-ended self-descriptive questions, the painting tradition of self-portrait in art therapy, and computer-based body image assessment tools. *Methods:* We engaged 11 university



students to create self-portraits using the Midjourney artist-software. Command prompts forming the basis of each image were subjected to category analysis, and semi-structured interviews were conducted with the participants. *Results:* Through statistical analysis, we identified significant individual differences in the self-descriptive categories mentioned by participants and their temporal variations. Future research will explore potential situational or intrapersonal influences behind these disparities, such as individual sensitivities to the flow state and the role of ego identity status. Feedback from the interviews revealed that participants perceived the process as an enlightening journey of self-discovery, experimenting with various facets of their identity. Many were astounded by the capabilities of the software and their own artistic creativity, although miscommunications in human-machine interactions occasionally hindered the creative process. *Conclusions:* The method we devised holds the potential to evolve into a valid and reliable tool not only for measuring self-image but also for assessing any attitude. Experiencing a sense of self-efficacy and the feeling of self-discovery could, in the long run, lay the foundation for its application in art therapy.

KEYWORDS

artificial intelligence, art, self-image, self-description, ego identity status, methodology

Open Access nyilatkozat. A cikk a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>) feltételei szerint publikált Open Access közlemény, melynek szellemében a cikk bármilyen médiumban szabadon felhasználható, megosztható és újraközölhető, feltéve, hogy az eredeti szerző és a közlés helye, illetve a CC License linkje és az esetlegesen végrehajtott módosítások feltüntetésre kerülnek. (SID_1)

