

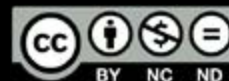
# Emanuelle Arielli és Lev Manovich fordította Pál Gyöngyi

ORCID: 0000-0002-7807-0557

**MATE KC Rippl-Rónai Művészeti Intézet**

[pal.gyongyi.katalin@uni-mate.hu](mailto:pal.gyongyi.katalin@uni-mate.hu)

Artcadia, ú.f. 2 (1), 53–63. (2023)



A cikkre a Creative Commons 4.0 standard licenc alábbi típusa vonatkozik: CC-BY-NC-ND-4.0.

## AI-aesthetics and the Anthropocentric Myth of Creativity

Suppose human creativity could be potentially replicated by mechanical processes. In that case, we would face a crossroads: either we could give up using the concept of creativity altogether, or if we hold to our common understanding of what creativity is, we could agree to apply this concept to non-human phenomena as well, as world champion Lee Sedol did when judging the performance of AlphaGo. However, the idea that artificial creativity discloses the mechanic nature of human creativity should also be met with a bit of critical detachment, particularly if we consider the specific case of the arts. In fact, artificial reproductions of human artifacts do not follow the same processes with which humans actually produced those artifacts. Nobody thinks that Mondrian followed procedures similar to the algorithm used in 1966 that generated pseudo-Mondrian, even though the public appreciated the artificial images more than the original ones...

## MI-esztétika és a kreativitás antropocentrikus mítosza

Tegyük fel, hogy mechanikus eljárásokkal potenciálisan reprodukálható az emberi kreativitás. Ebben az esetben válaszút elé kerülnénk: vagy teljesen lemondunk a kreativitás fogalmának használatáról, vagy ha ragaszkodunk a közmegegyezés szerinti kreativitás fogalmához, akkor meg kellene egyeznünk abban is, hogy ezt a fogalmat a nem emberi jelenségekre is alkalmazzuk, ahogyan Lee Sedol világbajnok tette az AlphaGo teljesítményének megítélésekor. Ugyanakkor szintén némi kritikai távolságtartással kell fogadnunk azt az elképzelést is, hogy a mesterséges kreativitás feltárja az emberi kreativitás mechanikus természetét, különösen, ha a művészetek sajátos esetét vizsgáljuk. Az emberi műalkotások mesterséges reprodukciói ugyanis nem ugyanazokat a folyamatokat követik, amelyekkel az emberek ténylegesen létrehozták ezeket a műalkotásokat. Senki sem gondolja, hogy Mondrian hasonló eljárásokat követett, mint az 1966-ban használt algoritmus, amely pseudo-Mondriant generált, annak ellenére, hogy a közönség jobban értékelte a mesterséges képeket, mint az eredetieket...

<http://manovich.net/index.php/projects/ai-aesthetics-and-the-anthropocentric-myth-of-creativity>

## Mesterséges esztétika

A 21. század eleje óta a számítástechnika, az adatelemzés és a mesterséges intelligencia fokozatosan belépett az esztétika területére: használják művészetet, zenét, könyveket és filmeket ajánló rendszerekben, illetve képek és videók automatikus szerkesztésénél. A mesterséges intelligenciát egyre gyakrabban használják új, mesterségesen generált művek, köztük műalkotások, zenék, design termékek és szövegek létrehozására is. 2016-ban például kiképeztek egy deep learning algoritmust Rembrandt 346 ismert festményének elemzésével arra, hogy megtanulja a festő stílusát, majd arra kérték, hogy hozzon létre egy vadonatúj portrét. Az eredmény kísértetiesen hasonlított egy valódi Rembrandt-festményre. Ugyanebben az évben a párizsi Sony Computer Science Laboratories kutatói kifejlesztettek egy DeepBach nevű neurális hálózatot, amely kóruskantátákat komponált J. S. Bach stílusában.<sup>1</sup> Azóta más, zenét generáló algoritmusok is születtek – ma már a Youtuberek zenei „Turing-tesztekre” hívják ki a nézőket, hogy képesek-e megkülönböztetni a mesterséges intelligenciával létrehozott zenei kompozíciókat az emberi kompozícióktól. A zenei képzettséggel rendelkező emberek számára a feladat egyszerűnek tűnik, de a naiv hallgatók számára ez nem mindig egyértelmű.<sup>2</sup> 2019-ben egy mesterséges intelligencia egy új okostelefon-modell számítási teljesítményét használta fel arra, hogy befejezze Schubert „Befejezetlen szimfóniáját” (1822./ 8. sz.).<sup>3</sup> Ebben a projektben azonban egy zeneszerző is közreműködött, aki kiválogatta és kiválasztotta az AI által generált legjobb dallamokat. Ugyanebben az évben

1 <https://arxiv.org/abs/1612.01010>

2 <https://www.youtube.com/watch?v=Pm-L31mVx0XA>; <https://www.youtube.com/watch?v=lv9W7qrYhbk>

3 <https://www.classicfm.com/composers/schubert/unfinished-symphony-completed-by-ai/>

a Deutsche Telekom egy nemzetközi zenei és mesterséges intelligencia-szakértőkből álló csapatot szervezett, hogy befejezzék Beethoven befejezetlen 10. szimfóniáját, és ezzel ünnepezzék meg a szerző születésének 250. évfordulóját. Az elkészült „Beethoven X – The AI Project” szimfónia premierje 2021. október 9-én volt Bonnban. Ezekben a példákban a számítógépeket már előre meglévő stílusokkal táplálják be, és cserében a gép ezeknek a stílusoknak megfelelő változatokat generál, megpróbálva némi újítást bevezetni. Nem teljesen új dalok vagy stílusok jönnek létre; ehelyett úgy tűnik, hogy ezek olyan példák, amit számítógépes *manierizmusnak* nevezhetünk.

Amikor egy gép megfest egy Rembrandtot, megkomponál egy Bach-szonátát vagy befejez egy Beethoven-szimfóniát, azt mondjuk, hogy ez nem eredeti vagy valódi *művészet*, hanem egyszerűen az emberi kultúra meglévő termékeinek komplex utánzása és reprodukálása. Ezáltal szembesülünk a kreativitás természetével kapcsolatos régi kérdéssel: milyen ötletek újrakombinálását, szokatlan analógiákat és fogalmi kapcsolatokat tekintünk az eredetiség jegyének? Kinek tulajdonítsuk a szerzőséget, ha egy műtárgy vagy kép elkészítéséhez olyan eszközöket, algoritmusokat és technológiákat használunk, amelyek egy művész vagy tervező intenciója szerint generálnak új művet és újraértelmezik azt? Mivel a termelési láncban egyre inkább megjelennek harmadik féltől származó szoftverek és a műveket egyre összetettebb beavatkozások alakítják (mint a fotó- és videoeffektek és szűrők vagy a retusáló algoritmusok), hogyan határozhatjuk meg, hogy az alkotás folyamatának melyik pontján történt kreatív innováció, és annak ki a szerzője? Mario Klingemann művész, az AI-art egyik

úttörője szerint: „Ha hallanál valakit zongorázni, megkérdeznéd: „A zongora a művész?” Nem, tehát ugyanez a helyzet itt is. Attól, hogy ez egy bonyolult mechanizmus, még nem változnak a szerepek”. Ebből a szempontból a mesterséges intelligencia művészetben való alkalmazása a *kiterjesztett esztétika* egyszerű példája lenne csak, amely új, látszólag kifinomultabb eszközöket használ az emberi művész szerzőségi irányításával. Ebben az esetben a mesterséges rendszer a művész és a programozó eszközének tekinthető, egy kifinomult eszköz, amelyet az alkotás folyamata során felhasználnak. Ugyanakkor továbbra is lenyűgöz bennünket az a gondolat, hogy az esztétika területén is tanúi lehetünk az autonóm mesterséges kreativitás megjelenésének, ragaszkodva a valódi mesterséges intelligencia eredeti elképzeléséhez, mint az autonóm és intelligens viselkedés megnyilvánulásához.

## Az esztétikai „Turing-tesztekről”, illetve: Mit is várunk egyáltalán az „esztétikát” gyártó gépektől?

2020-ban a Princeton Egyetem egyik végzős egyetemi hallgatójának projektje során egy úgynevezett Generatív Adverzális Hálózatot (GAN) használt arra, hogy olyan hagyományos kínai tájképeket készítsen, amelyek képesek becsapni az embert egy vizuális Turing-teszt során.<sup>4</sup> Az Alan Turing (1912-1954) által kidolgozott Turing-teszt eredeti megfogalmazásában ez egy olyan kritérium, amely alapján megállapítható, hogy egy mesterséges rendszer elérte az emberhez hasonló intelligenciát. Mi azonban nem mondanánk, hogy a princetoni diák által kifejlesztett GAN elérte az emberi szintű intelligenciát; ez

csupán egy olyan program, amely elég kifinomult ahhoz, hogy emberinek tűnő képeket generáljon. Ez tovább növeli a fogalmi zavart ebben a vitában.

Egyrészt az olyan fogalmak, mint az „intelligens” vagy a „kreatív” intuitívnak és egyszerűnek tűnnek, azt feltételezik, hogy mindenki képes felismerni az intelligens vagy kreatív viselkedés megnyilvánulását. Másrészt, amikor megpróbálunk egy működőképes és használható definíciót adni ezekre a fogalmakra, azt látjuk, hogy mennyire megfoghatatlanok. Ez a kérdés szembeállítja Alan Turingot Ludwig Wittgensteinnel (1889-1951), aki úgy vélte, hogy először nyelvi és fogalmi berögződéseinket kell tisztáznunk, amikor meg akarjuk érteni, mit értünk az olyan fogalmak alatt, mint az „intelligencia”. Turing 1939-ben részt vett Wittgenstein matematika-filozófiai előadásain, és ez utóbbi minden bizonnyal tisztában volt Turing tézisével a mechanikus gondolkodásról. Érdekes módon Wittgenstein véleményét olyan passzusokban fejti ki, mint az alábbi, amely a *Filozófiai vizsgálódások* (1953) című művéből származik:

„Képes lenne egy gép gondolkodni? – Lehetnének-e fájdalmai? – Nos, hívjuk tán az emberi testet ilyen gépnek? Hiszen ahhoz jár a legközelebb, hogy egy ilyen gép legyen. De hát egy gép nem tud gondolkodni! – Tapasztalati állítás ez? Nem. Csakis az emberről, és arról, ami hozzá hasonló, mondunk olyat, hogy gondolkodik. Mondjuk még babákról, és persze szellemekről is. Tekintsd a „gondolkodni” szót szerszámnak!”<sup>5</sup> (Wittgenstein, 1953: §359-360.)

5 Ludwig Wittgenstein, *Filozófiai vizsgálódások*, Budapest, Atlantisz Kiadó, 1998, 169. o. Fordította: Neumer Katalin.

4 <https://arxiv.org/pdf/2011.05552.pdf>

Wittgenstein szempontjából, mivel a szavak eszközök, fel kell tennünk magunknak a kérdést – ha egyáltalán szóba jöhet –, hogy milyen feltételekkel tudnánk használni olyan fogalmakat, mint a „gondolkodás” (vagy az „intelligencia” és a „kreativitás”) nem emberi, mesterséges entitások leírására.

A Turing-teszt egy olyan módszer, amellyel ellenőrizhető, hogy egy számítógépes interfészen keresztül beszélő gép, embernek minősül-e. Ezért a teszt az emberi viselkedés *mimikrijét* tekinti az intelligencia mutatójának, és elsősorban a verbális jelzésekre és a párbeszéd létrehozására alapul. Egyrészt Turing kritériuma ésszerűnek tűnik: ha valami nem különböztethető meg egy beszélgetés során az embertől, miért ne tulajdoníthatnánk neki intelligenciát? Másrészt azonban az emberek vonakodnak attól, hogy nem emberi entitásoknak könnyen megadják az intelligencia jelzőjét. Régebben úgy gondolták, hogy egy olyan gép, amely képes legyőzni egy sakk nagymestert, az valódi mesterséges intelligenciának bizonyulna.

Ez 1997-ben be is következett, amikor a DeepBlue legyőzte a világbajnok Garri Kaszparovot. Ekkor a sakkot pusztán kombinatorikai és számítási játékként kezdték el definiálni, nem pedig az intelligencia valódi tesztjeként; a mérföldkövet pedig más, összetettebbnek tartott és inkább kreatív intuíciókon alapuló játékokra, mint a Go-ra helyezték át. 2016-ban azonban a Google AlphaGo-ja legyőzte a világbajnok Lee Sedolt (sz. 1983), mégsem érezzük továbbra sem úgy, hogy ekkor „igazi” intelligenciát sikerült teremteni. Vagy gondoljunk a chatbotokra. Turing 1950-es tanulmánya szerint<sup>6</sup> a század végére a gépek öt perc beszélgetés után az emberek egyharmadát képesek lesz-

nek becsapni. 2014-ben a bírálók 33%-a tartotta embernek az „Eugene Goostman” nevű chatbotot, ami gyakorlatilag átment Turing tesztjén (itt meg kell jegyezni, hogy Goostmant úgy programozták, hogy egy 13 éves ukrain odesszai tinédzser szeszélyességét és furcsaságait szimulálja).

Minden egyes alkalommal, amikor egy technológiai mérföldkövet elérünk, úgy tűnik, hogy a céltábla egyre távolabb kerül. Wittgenstein szempontjából ennek oka nem abban rejlik, hogy az újabb technológiai mérföldkövek nem elég meggyőzőek ahhoz, hogy biztosítsanak minket arról, hogy valódi intelligenciával van dolgunk. A kérdés valójában egyáltalán nem empirikus, hanem azokkal a feltételezésekkel kapcsolatos, amelyeket az olyan fogalmak használata és besorolása során teszünk, mint az intelligencia és a kreativitás. Ez vezet el az úgynevezett **Tesler-tételhez**, amely kimondja, hogy: A mesterséges intelligencia az, amit még nem csináltak meg (vagy fordítva, az intelligencia az, amit a gépek még nem értek el).<sup>7</sup> Ma egy olyan alkalmazás, mint a Siri, képes emberhez hasonló párbeszédet folytatni. Az Open-AI által nemrégiben kifejlesztett GPT-3-on alapuló szöveggenerátor – amelyet 570 GB internetes szövegekből álló adathalmazzal képeztek ki – képes olyan kifinomult újságcikkeket írni, amelyek megkülönböztethetetlenek az ember által generált szövegektől. Azonban éppen azért, mert tudjuk, hogy ezek egy kifinomult programozás termékei, még mindig úgy gondoljuk, hogy nem valódi intelligenciáról van szó, nemhogy szándékosságot vagy tudatosságot tulajdonítanánk ezeknek a rendszereknek.

Másképp fogalmazva, nem vagyunk hajlandók az „intelligencia” szót használni ilyen esetben; általában akkor használ-

6 E definíció szerzője Larry Tesler, egy jól ismert informatikus, aki a Xerox PARC-nál, az Apple-nél és az Amazonnál dolgozott.

6 Alan M. Turing, „Computing Machinery and Intelligence” (Számítógépek és intelligencia). Mind, 1950, 59., 433-460. o.

jük, amikor személyekre utalunk, és mint Wittgenstein mondta, a szavak olyan eszközök, amelyeknek sajátos használatához hozzá vagyunk szokva. Ezért Tesler tételének további következménye, hogy a „mesterséges intelligencia” kifejezés minden használata, például az arcfelismerés, a spamszűrők, a számítógépes látás, a beszédgenerálás stb. kontextusában, definíció szerint *nem* mesterséges intelligencia, hanem olyan technológia, amely komplex optimalizációs algoritmusokat használ. Csak marketing okokból nevezik „MI”-nek.

Ha az intelligencia tulajdonítása egy olyan horizontvonal, amelyet soha nem lehet elérni, akkor elgondolkozhatunk azon, hogy vannak-e egyáltalán emberi képességek, amelyek e vonalon túl találhatóak: minden alkalommal, amikor a gépek „megoldanak” egy bizonyos emberi képességet, ez a képesség megszűnik valódi intelligenciának lenni, és mechanikusabbnak bizonyul, mint amilyennek látszott. Ennek azonban következményei lehetnek magáról az emberi intelligenciáról alkotott elképzeléseinkre is.

A művészetek és a művészek itt jönnek a képbe. A mesterséges intelligencia és az esztétika találkozása azért kulcsfontosságú, mert a művészetet eredendően emberi területnek tekintik, és annak bonyolultsága és összetettsége már régóta úgy tűnik, hogy nem alkalmas algoritmikus redukcióra. Sokan a művészetet, az esztétikát és a kreativitást tartják az emberi képességek csúcának; ezért úgy tekintenek rájuk, mint a mesterséges intelligencia előretörésével szembeni utolsó barikádra, amely messze van attól, amit a technológiai fejlődés reprodukálni képes. Ha a Turing-teszt hagyományos definíciójánál maradunk, az esztétika területén ez annak feleltethető meg,

ha olyan műalkotást (legyen az szöveg, párbeszéd vagy műalkotás) tudunk gépekkel létrehozni, amely képes becsapni egy embert. De miért kellene viszonyítási alapnak tekinteni azt, hogy hasonlítson egy ember által létrehozott műalkotáshoz? Mi a helyzet az olyan innovatív, szép vagy lenyűgöző tervekkel vagy művészeti formákkal, amelyek egyértelműen *nem emberinek* tűnnek? Egy olyan Turing-teszt, amelynek célja a megfigyelő megtévesztése, ebben az esetben alkalmatlan lenne.

Ezért érdemes lenne a Turing-teszt célját az eredetileg alapul szolgáló egyszerű „utánzó játékon” túlmenően felülvizsgálni, és másképp meghatározni. Például azt mondhatnánk, hogy egy gép akkor megy át egy ilyen teszten, ha a következő feltételek bármelyike teljesül:

- 1) *Kiemelkedő* emberi teljesítményt ér el (azaz olyasmit hoz létre, ami kiemelkedő a szépség, kellemesség, „bámulatosság” stb. tekintetében), tekintet nélkül arra, hogy hasonlít-e az emberi kulturális viselkedésre.
- 2) Képes a *kreativitásra*, vagyis újítani tud.
- 3) *Autonóm* módon működik, azaz a gép látszólag képes valami váratlan, a programozók kezdeti paramétereitől és bemeneteitől eltérő dolgot produkálni.

A mesterséges intelligencia (1) kiváló teljesítményének hírhedt példája, hogy a programok legyőzik az embert olyan játékokban, mint a sakk vagy a Go. De még az esztétikában sem újdonság az a képesség, hogy az emberi képességeken túlmutató dolgokat tudnak előállítani: már 1966-ban egy algoritmus olyan Mondrian-festményeket generált, amiket a közönség esztétikailag kellemesebbnek ítélt, mint a valódi

8 Noll, „M. Ember vagy gép: Piet Mondrian „composition with lines” (1917) és egy számítógép által generált kép szubjektív összehasonlítása”. The Psychological Record, 1966, 16., 1-10. o.

9 Lásd az olyan platformokat, mint az aiva.ai, amelyek lehetővé teszik új, szerzői jogtól mentes zenék generálását a már létező dalok stílusát követve.

10 <https://arxiv.org/pdf/1410.6142v3.pdf>

Mondrian-vásznakat.<sup>8</sup> Egy olyan forgatókönyvre is gondolhatunk, amelyben a mesterséges rendszerek jobb zenét, jobb könyveket, meggyőzőbb forgatókönyveket fognak előállítani, nem feltétlenül a műkritikusok, hanem egyszerűen a kulturális ipar szemszögéből: azaz olyan alkotásokat, amelyek nagy közönség- és kereskedelmi sikert aratnak. A dallamokat vagy dalszövegeket generáló algoritmusok (vagy Mondrian vagy más híres művész stílusában festő algoritmusok) a költség/bevétel arányt figyelembe véve is felülmúlják az emberi produkciót, ha más nem, akkor tisztán gazdasági szempontból, hiszen egy művész zenei vagy festői stílusára nem létezik védjegyoltalom.<sup>9</sup>

Ami a kreativitást illeti, ez önmagában is egy megfoghatatlan fogalom, és hosszú viták tárgya a filozófiában és a kognitív tudományokban. A „kreativitás Turing-tesztjét” más néven Ada Lovelace-tesztnak is nevezik, Ada Lovelace 19. századi matematikusnak a kreatív gépek lehetőségéről tett megjegyzései alapján. Egy ilyen tesztben megmutatnánk egy gép által generált műalkotást, és megkérnénk a közönséget, hogy ítélje meg, hogy kreatív-e (és milyen mértékben az).<sup>10</sup> A kreativitás és az újdonság megítélése részben szubjektív kérdés, és gyakran attól függ, hogy mi, emberek miként *tulajdonítunk* kreativitást egy viselkedésnek. Az egyik szűk értelmezés például azt feltételezi, hogy csak az ember lehet képes kreativitásra, és hogy csak akkor beszélhetünk kreatív viselkedésről, ha az ember *öntudatos* és tudatában van annak, amit csinál. Ugyanakkor gyakran használjuk ezt a fogalmat liberálisabb és metaforikusabb értelemben is, amikor például azt mondjuk, hogy „a természet kreatív” (például egy új organizmus vagy egy új vírus létrehozásában). Ebben az esetben csupán a kre-

ativitás fogalmát alkalmazzuk egy olyan jelenségre, amely *váratlan*, azaz tudomásunk szerint korábban nem létezett.

Ebből a szempontból minden olyan véletlenszerű és meglepő folyamatot, amely nem könnyen megjósolható, kreatív-nak kell tekinteni; nem véletlen, hogy a 20. századi avantgárd művészek, mint például a dadaisták, sztochasztikus folyamatokkal kísérleteztek. A véletlenszerű folyamatok önmagukban azonban nem elegendőek ahhoz, hogy valamit kreatív-nak nevezzünk: elvárjuk, hogy valami kreatív-nak értelme is legyen, például egy régi probléma újszerű megoldása vagy valamilyen feladat vagy probléma elvégzésének jobb módja.

A kreativitás meghatározásának kihívásaihoz hasonlóan az *autonómia* meghatározása sem könnyű. Egy gép akkor tűnik autonómnak, ha az eredeti programozásától független viselkedést mutat – vagyis ismét csak akkor, ha a megfigyelő számára váratlanul és kiszámíthatatlanul viselkedik. Egyrészt nincs egyértelmű kritérium az autonómiára: autonóm-e egy egysejtű organizmus? Mi a helyzet a rovarokkal? Az autonómia megítélésében is nagy mértékben szerepet játszik a szubjektivitás.

## A mesterséges intelligencia mint az emberi képességek kritikus tükré

Ludwig Wittgenstein filozófus, aki a számítás és a gondolkodás gépesítésének lehetőségéről vitázott Alan Turinggal, más értelmezést kínált ez utóbbi híres tesztjéről. Wittgenstein szerint a teszt nem egy olyan módszer, amellyel azt vizsgáljuk, hogy egy gép képes-e becsapni egy megfigyelőt és embernek kiadni magát. A teszt ehelyett azt mutatná meg szerinte, hogy az



**ember milyen mértékben viselkedik és gondolkodik gépiesen.** Ha ebből a szempontból nézzük a dolgokat, akkor kijózanító az emberi kreativitást szimuláló alkalmazások kifejlesztése. Például egy olyan program, amely képes fülbemászó dallamokat vagy lenyűgöző forgatókönyveket generálni, megmutatja, hogy mennyire „mechanikusak” azok a folyamatok, amelyeket egyébként intuitívnak és szabadnak tartunk. Ha ebből a szögből néznénk, akkor a Turing-teszt céljától függetlenül a tesztet sikeresen teljesítő gépek arra mutatnának rá, hogy az emberek sokkal mechanikusabbak, mint gondolnánk. Ennek eredményeképpen talán túlértékeljük a kreativitást mint emberi képességet, egyszerűen azért, mert nem értjük a működését. Az a tény, hogy bizonyos emberi folyamatok mechanikusabbnak és procedurálisabbnak tűnnek, mint azt feltételezzük, megkérdőjelezi a kreatív intuíciónál alkotott tipikusan romantikus felfogásunkat. Nem szabad elfelejteni, hogy a tiszta kreativitás eszméje az egyén autonómiájának felmagasztalásából ered, amely csak a modernitásban honosodott meg. Ez nem volt elképzelhető az ókorban, ahol az uralkodó nézet szerint az ember csak arra volt képes, hogy emlékezzen (a platóni *anamnézis* értelmében), rekonstruáljon és reprodukáljon már létező dolgokat. A művész ebben az értelemben felfedező volt, nem pedig alkotó; a művészet a valóság ügyes utánzása, nem a pusztá feltalálás, hanem a mesterség volt. Az igazi kreativitás, az ókori és középkori értelemben vett creatio (ex-nihilo), kizárólag az isteni hatalom előjoga volt.<sup>11</sup>

A művészeti stílusok történelmi fejlődését hagyományosan kiszámíthatatlan kreatív ugrások következményének tekintjük, amelyeket rekonstruálhatunk utólag, de előre megjósolni nem

tudunk. Az evolúciós algoritmusok egyes alkalmazásai azonban úgy tűnik, hogy másra engednek következtetni. A vizuális művészetekkel kapcsolatban például Lisi és munkatársai (2020)<sup>12</sup> megmutatták, hogy lehetséges előrejelezni a képzőművészetekben a stílusfejlődést egy olyan rendszer betanításával, amely nagy képadatbázisok elemzésével extrapolálja bizonyos evolúciós törvényszerűségeket, és képes az időben egymást követő új stílusok szerinti képeket generálni. A szerzők szerint a rendszer meglepő módon olyan előrejelzéseket generált, amelyek nagymértékben tükrözik a képzőművészet történetében létrejött stílusok tényleges fejlődését, kiemelve ezzel egyes stílusfejlődések „algoritmikus” jellegét. Ez azt jelenti, hogy ez a fejlődés nem történelmi véletlenek vagy egyedi művészek spontán találmányainak köszönhető, hanem belső formai törvényszerűségek szinte szükségszerű egymástánjából adódik.<sup>13</sup> Egy ilyen rendszer ráadásul képes lehet a vizuális művészet *jövőbeli* stílusainak előrejelzésére is. Ezeknek a fejleményeknek nem kell determinisztikusnak lenniük, de mindazonáltal az adatelemző rendszerek véges kombinációkból kikövetkeztetett és reprodukált jelenségének tekinthetjük.

Ezek a példák arra engednek következtetni, hogy a „kreativitás” egy olyan címke, amelynek olyan jelenségeket tulajdonítunk, amelyek mögöttes folyamataival nem vagyunk tisztában. Amikor például 2016-ban Lee Sedol, a Go világbajnokát legyőzte az AlphaGo, akkor azt nyilatkozta, hogy a program hihetetlenül kreatív lépésekre képes, és ezzel felfedte, hogy bizonyos lépések vagy játéktaktikák, amelyeket az emberek kreatívnak gondoltak, valójában meglehetősen kiszámíthatóak. A kihívás második játszma során az AlphaGo olyan

12 Lisi E, Malekzadeh M, Haddadi H, Lau FD-H, Flaxman S. „Modelling and forecasting art movements with CGANs.”, 2020, Royal Soc. Open Sci. 7: 191569. <http://dx.doi.org/10.1098/rsos.191569>.

13 Magának a formának a belső logikájára vonatkozó hasonló gondolatot vetett fel George Kubler *The Shape of Time* című 1962-es munkája is.

11 Tatarkiewicz, Władysław, *A History of Six Ideas: an Essay in Aesthetics*, 1980, The Hague: Martinus Nijhoff.

lépést tett (37. lépés), amelyet sok kommentátor szokatlanul kreatívnek minősített, és váratlanul érte a játékost, így tudott végül a számítógép győzni. Az, hogy ezt a konkrét lépést a megfigyelők kreatívnek ítélték, az betudható annak, hogy a játékosok és a szakértők nem voltak tisztában azzal, hogy mi az AlphaGo mögöttes stratégiája. A gép szemszögéből nézve valójában ez a lépés egy olyan kiértékelés eredménye volt, amely ugyanazokat az optimalizációs folyamatokat követte, amelyekkel a rendszer minden másik lépést is kiválasztott. Ebben a tekintetben valamit kreatívnek nevezni gyakran a megértésünk hiányának a fokmérője: amit ismerünk, az hétköznapiak számítt, amit viszont nem ismerünk, azt rendkívüli dolognak tartjuk. Más szóval, ha úgy gondoljuk, hogy az emberek kreatívak, az MI-k pedig nem, az azért van, mert jobban értjük, hogyan működik az MI, míg az emberek működését még mindig nem értjük eléggé. A technológiai fejlődés gyakran nyilvánvalóvá teszi, hogy a rendkívülinek tartott jelenségek valójában hétköznapi folyamatok eredményei.<sup>14</sup>

14 A kreativitás „a hétköznapi folyamatok rendkívüli eredményeiben” áll. R.J. Sternberg, T.I. Lubart, „Investing in creativity” American Psychologist, 1996, 51, 681. o.

## Nincs szellem, csak egy héj?

Tegyük fel, hogy az emberi kreativitás potenciálisan megismételhető mechanikus folyamatokból áll. Ebben az esetben válaszút elé kerülnénk: vagy teljesen lemondunk a kreativitás fogalmának használatáról, vagy ha ragaszkodunk a kreativitás fogalmának közös értelmezéséhez, akkor megegyezhetünk abban, hogy ezt a fogalmat a nem emberi jelenségekre is alkalmazzuk, ahogyan Lee Sedol világbajnok tette az AlphaGo teljesítményének megítélésakor.

Ugyanakkor azt az elképzelést, hogy a mesterséges krea-

tivitás feltárja az emberi kreativitás mechanikus természetét, szintén némi kritikai éllel kell fogadnunk, különösen, ha a művészetek sajátos esetét tárgyaljuk. Az emberi műalkotások mesterséges reprodukciói ugyanis nem ugyanazokat a folyamatokat követik, amelyekkel az emberek ténylegesen létrehozták ezeket a műalkotásokat. Senki sem gondolja, hogy Mondrian hasonló eljárásokat követett, mint az 1966-ban használt algoritmus, amely pseudo-Mondriant generált, még akkor sem, ha a közönség jobban értékelte a mesterséges képeket, mint az eredetieket. Nem hagyhatjuk figyelmen kívül a festő stilisztikai újítása mögött rejlő szimbolikus, történelmi és konceptuális jelentéseket, sem a művésznek a festészet fejlődésében, különösen az absztrakció, a figuratív művészet, az expresszionizmus és a minimalizmus terén betöltött szerepét. Más szóval, az algoritmus nem reprodukálta azt a **kulturális folyamatot**, amelyen keresztül Piet Mondrian eljutott absztrakciójához.

Ehelyett a programozók csak formai szinten utánozták a végterméket. Mondrian festményeit úgy csodáljuk, mint a művésznek az elkészítésükhöz vezető útjának végső kifejeződését, a festészet történetében betöltött kulturális szerepét. E tényezők nélkül festményeit csak művészi érték nélküli érdekes geometriai mintázatoknak látnánk. Hasonlóképpen, Lucio Fontana vágott vászna is csak egy vágással ellátott vászon lenne, amelyet könnyen reprodukálhatnánk egy program vezérlésének a segítségével és egy késsel felszerelt mechanikus karral (mint amilyeneket a robotsebészetben már használnak). Az ilyen művek előállításának egyszerűsége megmutatja, hogy többről van szó a látszatonál, valójában ez rámutat



arra a különbségre, ami a kortárs művészetre jellemző esztétikai és a művészi érték között feszül. E művek esztétikai értékelése során történelmi, konceptuális és szimbolikus dimenziót látunk a tárgyban, és a vászon formális felületén látható dolgokon túl konkrét szándékokat tulajdonítunk az alkotónak. Szimbolikus jelentések, affektív asszociációk és kulturális utalások sokasága gazdagítja a műtárgyat; ennek a felismerésére csak akkor vagyunk hajlandóak, ha olyan alanytól származnak a művek, akinek teljes joggal tudatos szándékot tudunk tulajdonítani. Ezzel szemben vonakodunk jelentőséget ráruházni arra, amit egy algoritmus állít elő, mert lélek nélkülinek tartjuk.

## Technoanimizmus és a Pygmalion-effektus

A jelenségeknek szándékosságot tulajdonító természetes hajlamunk az, ami lehetővé teszi, hogy egy gépet intelligensnek vagy akár tudatosnak ismerjünk el. A gyerekek is így tesznek a játékokkal és más tárgyakkal; néha a felnőttek is emberhez hasonló viselkedést tulajdonítanak például növényeknek vagy kisállatoknak. Sok jelenlegi és múltbeli kultúra mélyen animista azokkal a természeti jelenségekkel szemben, amelyeket nem tudtak megmagyarázni kauzális okokkal és a fizika törvényeivel. Ezekben a világnézetekben a valóságot gazdagon benépe-sítik a nem emberi ágensek, legyenek azok növények, állatok, meteorológiai vagy földrajzi jelenségek. Hogyan tudná egy kőkorszakból származó ember értelmezni a mai automata ajtókat viselkedését, amelyek minden alkalommal kinyílnak, amikor valaki eléjük lép? Valószínűleg azt gondolná, hogy intelligenciával

és céllal rendelkeznek. Naivitás lenne ezeket az animista nézeteket egyszerűen tévesnek minősíteni: jobb magyarázat híján a szándékot tulajdonító modellek gyakran jó magyarázóerővel bírnak az ilyen jelenségek leírásában. Az őskori férfi vagy nő értelmezésében az tűnhet természetesnek, hogy az ajtó magától ki **akar** nyílni, és át akarja engedni az embert. Hasonlóképpen, a mesterséges intelligenciáról alkotott felfogásunk erősen függ attól, hogy mit vetítünk ki rá és intenciót tulajdonítunk-e a mesterséges, nem emberi entitásoknak.

Bár az a hajlam, hogy lélekkel ruházzuk fel a nem emberi lényeket, kulturális háttérünktől, vallási érzékenységünktől és egyéni meggyőződésünktől függ, manapság az az uralkodó feltevés, hogy csak az emberek (és kisebb mértékben egyes állatok) rendelkeznek valódi szándékossággal és tudatos működésmóddal. Amikor szándékosságot tulajdonítunk más entitásoknak (egy ajtónak, egy játéknak, egy virtuális asszisztensnek, az időjárásnak), azt mondjuk, hogy ezt mindig csak **metaforikus** értelemben tesszük, egyfajta fiktív hozzáállásként, amelyben úgy viselkedünk, „mintha” az entitásnak lenne valamilyen tudatossága, de anélkül, hogy ezt valóban elhinnénk. Hasonló történik akkor is, amikor egy film vagy egy regény szereplőit úgy vesszük, „mintha” valóságosak lennének, még akkor is, ha tudjuk, hogy nem azok.<sup>15</sup> Meg kell azonban jegyezni, hogy a határ képlékeny a valódi és a látszólagos intencionalitás érzékelése között. Például az olyan háziállatokat, mint a macskák és a kutyák, valódi tudattal rendelkezőnek tekintjük. Sokak számára ez a rovarokra vagy baktériumokra is érvényes, de van, aki ezzel már nem ért egyet. Mások ezzel szemben még a növényekbe is belelátanak személyiséget, míg

15 K. Walton, *Mimesis and make-believe*, 1990, Harvard University Press.

megint mások csak „mintha” módon teszik ezt. Az egyéni és kulturális különbségek határozzák meg, hol húzódik a határ a szándékosság valós és fiktív megítélése között.

Ami a technológiai eszközöket illeti, a velük szembeni „úgy-csinálás” birodalmában vagyunk: megtanulunk interakcióba lépni az olyan virtuális asszisztensekkel, mint Alexa, akik úgy beszélnek, „mintha” emberként hallgatnának minket. Ahogy ezeknek az eszközöknek a komplexitása és rugalmassága növekszik, elkezdhetjük úgy tekinteni, mintha teljes értékű, tudatossággal felruházott entitásoknak tekintenénk őket. Ha ez megtörténik, annak egyik oka kétségtelenül a technológiák fejlődése lesz. A másik ok azonban az előítéletek kulturális leküzdésének lesz a következménye: ma még mindig szívesebben adnánk szándékosságot egy rovarnak, mint Alexának, függetlenül attól, hogy az utóbbi komplexitása, a tudáshoz való hozzáférése és a velünk való interakcióra való képessége messze felülmúlja egy bogárét. Sőt, talán azt az elképzelést is meg kellene kérdőjeleznünk, hogy a (dolgokra, állatokra és nem emberi entitásokra alkalmazott) „mintha” intencionalitás csupán metaforikus kapcsolatban áll a „valódi” intencionalitással. Lehet, hogy éppen ellenkezőleg: az igazi intencionalitás szűk (csak az emberekre alkalmazott) felfogása ered valójában a „mintha” intencionalitásból, amely ahhoz a természetes és mélyen bennünk gyökerező hajlamunkhoz kötődik, hogy előszeretettel tulajdonítunk szándékot a jelenségek széles körének.<sup>16</sup>

Ebben a vitában néha két, látszólag ellentétes álláspontot figyelhetünk meg: az egyik csak az emberekben (és néhány állatban) látja a valódi szándékosságot, a másik a nem embe-

ri entitásoknak is tudatosságot tulajdonít, egyfajta naiv animizmus révén „humanizálva” őket. Mindkét álláspont azonban ugyanazt az antropomorf szemléletet vallja a tudatos működésről és az intencionalitásról, az egyik esetben megvonják, a másikban pedig ráruházzák a tudatosságot a nem emberi entitásokra. Egy alternatív nézet lenne, ha meghatároznánk a tudatosság egy speciális fogalmát a szubperszonális folyamatok, a nem-emberi entitások és a mechanikus jelenségek számára. Tehát nem arról lenne szó, hogy humanizálnánk azt, ami nem emberi, hanem arról, hogy megértenénk a nem emberi és nem antropocentrikus intenciót. Ebben a kérdésben a mesterséges intelligenciáról alkotott felfogásunk megváltozása a tudatos működésmód és a kreativitás antropocentrikus szemléletének meghaladását is eredményezné. Ez azt az irányt követné, amelyet már a klasszikus poszthumán elméletek, például Donna Haraway és Rosi Braidotti munkái előrevetítenek, vagy Bruno Latour javaslata, miszerint „újra kell gondolni a társadalmat” az emberi és nem emberi entitások figyelembevételével, amely nemcsak a nem emberi (állati vagy növényi), hanem a mesterséges teremtményeket is magában kellene foglalnia.

Hozzátehetnénk, hogy a kulturális javak előállításának bizonyos formáiban fontosnak tűnik a tudatos működés és a szándékosság kérdése, de másokban nem feltétlenül. Egy díszítőminta, egy bútordarab vagy egy autó nem igényel (mindig) szerzői mélységet; nem szükséges, hogy jelentéseket lássunk vagy a szerző gondolatairól okoskodjunk. Még egy esztétikailag magával ragadó fülbemászó dal is arra készíthet bennünket, hogy figyelmen kívül hagyjuk, hogy van vagy nincs

16 Ezt a gondolatot közismerten Daniel Dennet képviseli. Lásd D. Dennet, *The Intentional Stance*, 1987, MIT Press.

mögötte álló szerzői szándék. Hasonlóképpen, egy filmet is lehet pozitívan értékelni pusztán azért, mert önmagában is magával ragadó és szórakoztató, anélkül, hogy el kellene gondolkodnunk azon, hogy az író vagy a rendező mit akart mondani. A mesterséges intelligencia-művészet generálásánál így érdekes próbatétel lesz annak meghatározására, hogy mely területeken érezzük szükségét egy műalkotás mögött felismerhető tudat jelenlétét, és mely területeken tudunk megenni nélküle.

Egyrészt azt gondolhatnánk, hogy egy dal csak akkor lehet sikeres, ha kielégíti azt az igényünket, hogy szándékosságot tulajdonítsunk az írójának és előadójának, miután lehetővé teszi számunkra, hogy beeléssünk szimbolikus, érzelmi és személyes mélységet. Másrészt a virtuális popsztárok sikere a kelet-ázsiai kultúrákban (mint például Hatsune Miku és számos K-pop „avatár”-zenekar, amelyek közül némelyikük kizárólag mesterséges intelligencia által vezérelt) rámutat arra, hogy a közönség hogyan tud érzelmileg elköteleződni egy fiktív előadó mellett is, hogyan tudják a rajongók követni őket a közösségi profiljukon, elmenni a koncertjeikre, és megvásárolni az őket ábrázoló kütyüket.<sup>17</sup> Elmehetnénk odáig, hogy azt mondjuk: a rajongók nem annak ellenére szeretik őket, hogy felvállaltan hamisak, hanem valójában éppen azért, mert azok. Így egyetlen valódi személy sem áll az útjába annak az igényüknek, hogy vágyaikat és képzeletüket egy virtuális avatárra vetítsék. Mint Pygmalion mítoszában a művész, aki megvetette a valódi nőt, de beleszeretett az idealizált női szoborba, mi is mélységes kielégülést találhatunk abban, hogy egy olyan géppel foglalkozunk, amelynek lélektelensé-

ge lehetővé teszi számunkra, hogy megtöltsük a tökéletes lélekről alkotott elképzelésünkkel. Ezért képezhetünk töretlen átmenetet a „mintha” és a „valódi” lélek megítélése között: ugyanúgy érdekel bennünket egy virtuális személyiséggel való foglalkozás, mint ahogyan egy filmsorozat (vagy rajzfilm) szereplőjének drukkolunk, vagy egy regény főhősének megszállottjai vagyunk. Ugyanez vonatkozik az ezeket a dalokat vagy történeteket előállító művészre is: már nem aggódunk amiatt, hogy az alkotófolyamat algoritmus eredménye vagy emberi alkotás volt-e, amíg érzelmileg elkötelezettek vagyunk.

Fordította: Pál Gyöngyi

<sup>17</sup> <https://www.flyfm.audio/flycelebrity-the-future-of-k-pop-all-artificial-intelligence-ai-girl-group-eternity-drops-their-mv/>