

MESTERSÉGES INTELLIGENCIA-NARRATÍVÁK A TUDOMÁNYOS  
FANTASZTIKUMBAN ÉS AZ ÚJMÉDIÁBAN

**Szerző:**

Babos Orsolya  
Miskolci Egyetem

Szerző e-mail címe:  
babos.orsi@rocketmail.com

**Lektorok:**

Farkas Zsolt (PhD)  
Miskolci Egyetem

H. Nagy Péter (PhD)  
Selye János Egyetem (Szlovákia)

...és további két anonim lektor

**Absztrakt**

A tanulmány először röviden áttekinti a tudományos fantasztiikum különböző médiumain elterjedt mesterséges intelligencia-narratívákat, majd pedig három videójáték példáján keresztül vizsgálja meg az MI-vel kapcsolatban felmerülő etikai-filozófiai kérdéseket, kitérve az erős és a gyenge MI megkülönböztethetőségének a problémájára, és arra, hogy a médium különlegességét adó interaktív elemek milyen módon nyújthatnak számunkra gyakran az irodalom és a film erényeit egyesíteni képes befogadói élményt.

**Kulcsszavak:** mesterséges intelligencia, narratíva, tudományos fantasztiikum, videójáték

**Diszciplínák:** irodalomtudomány, médiatudomány

**Abstract**

*ARTIFICIAL INTELLIGENCE-NARRATIVES IN SCIENCE FICTION AND NEW MEDIA*

This study briefly reviews some of the popular artificial intelligence narratives in the science fiction genre and studies ethical and philosophical questions in relation to the A.I. through the analysis of three video games. This study highlights, that it is not easy the distinction between the strong and the weak A.I., and shows that how can special interactive elements of media give experiences that can synthesising the virtues of literature and film.

**Keywords:** artificial intelligence, narrative, science fiction, video games

**Disciplines:** literary studies, media studies

Babos Orsolya (2021): Mesterséges intelligencia-narratívák a tudományos fantasztikum-  
ban és az újmédiában. *Mesterséges intelligencia – interdiszciplináris folyóirat, III. évf. 2021/2.*  
szám. 55-76. doi: 10.35406/MI.2021.2.55

A mesterséges ember létrehozásával kapcsolatos történetek már az antikvitás kora óta jelen vannak az emberiség kollektív tudatában. Ilyenek a Taloshoz, vagy Galateiához kötődő mítoszok, vagy a gólemek, homunkuluszok megalkotására tett kísérletek, illetve a róluk szóló legendák. De az emberiség még soha nem volt ilyen közel ahhoz, hogy ezt az álmát megvalósulni lássa, mint napjainkban. A technológiai fejlődés következményeképp a mesterséges intelligencia a mindennapjaink részévé vált, elég például a GPS-ek navigációs rendszereire, vagy a közösségi oldalak képfelismerő algoritmusaira esetleg az okostelefonok személyi asszisztenseire gondolnunk.

Ezekből a példákban is látszik, hogy az, amit a közbeszédben MI-nek (mesterséges intelligenciának) nevezünk, egyelőre az élet speciális területeire kifejlesztett segédalkalmazások halmaza, ami még egyáltalán nem egyenlő az öntudatra ébredt géppel, vagy a mesterséges tudattal, melynek különböző archetípusait a tudományos fantasztikus szerzők alkották meg.

Egy saját, korábbi tanulmányomból idézve: „a Mary Shelley Frankensteinjében szereplő teremtmény is elfogható a robotos történetek hőseinek előzményeként és az istent játszó ember ártatlan áldozataként: gyermeki tisztaságú értelem förtelmes test-

be zárva. Eredete természetellenes, hibriditása pedig idegenné teszi a földi élőlények között. Az emberek félnek tőle, mert nem ismerik és mert a felszín alapján ítélnék, a teremtője pedig megtagadja tőle a saját fajtájából való társat és ezzel beteljesíti azt, amitől kezdettől fogva rettegettek, akik csak látták a lényt: a szörnyeteggé tett kreatúra haragját. Ha az 1800-as évek első felében még nem is lehet szó mesterséges intelligenciáról, a tudomány rossz célokra való felhasználásától, vagy a hübris bűnébe eséstől való félelem egyértelműen ott lapul a sorok között” (Babos, 2020, 1. o.).

Irving J. Good már 1965-ben felvetette egy ún. intelligenciारobbanás lehetőségét, amelyet ma már inkább szingularitásként emlegetnek, és mely során „ha egy ultra-intelligens gép képessé válna további intelligens gépek létrehozására, az hamarosan az embernél sokkal nagyobb intellektussal bíró mesterséges értelem megalkotásához vezetne” (Good, 1966).

A következőkben röviden összefoglalom, hogy milyen típusú MI-narratívákkal találkozhatunk a tudományos fantasztikum területén, ezután pedig három videojátékban fogom megvizsgálni a bennük megjelenő MI-k és a játékos karakter, valamint környezetük viszonyát, illetve a programok mesterséges és virtuális tudattal kapcsolatos problémafelvetéseit.

Hogy mi a haszna a fiktív MI-narratívákkal való foglalkozásnak? A sci-fi már régóta ez egyik legbefolyásosabb alakítója a jelen folyamatokról és a lehetséges jövőkről szóló elképzeléseknek, mindenféle névumok elméleti tesztelésének. Ilyen irányú vizsgálataink meghatározhatják, vagy adott esetben újra is konfigurálhatják az MI kapcsán kialakított koncepcióinkat. Előzetesen feltárhatjuk, hogy elméleti síkon milyen nehézségekkel kell szembenéznie azoknak, akik egy mesterséges, autonóm gondolkodó entitást szeretnének létrehozni. Megfogalmazhatjuk az erős MI kapcsán megjelenő várakozásainkat, félelmeinket erkölcsi aggályainkat; miként határozzuk meg a jogi státusát, még mielőtt megjelenne az első valódi MI; miféle korlátokat, normarendszert kell kidolgoznunk számára; lehet-e egy ennyire komplex dolog hibátlan; és ellenőrizhető-e egyáltalán, ha intelligensebb nálunk? Min, kin múlik, hogy jótékony vagy kártékony lesz-e?

### **Hogyan kategorizálhatjuk az MI-ket?**

Sok, főként újabb filmes és irodalmi MI narratívára jellemző, hogy testben élő MI-ket, azaz robotokat szerepeltetnek (v.ö.: Mező, Mező és Mező, 2019). Olykor pozitív, de leginkább negatív jövőképeket projektálnak és meglehetősen sablonosak szereplőiket és a történetek fő fordulatait tekintve. A disztópikus vonalon ilyen például a „gonosz” robotok lázadása (Terminátor), és az emberiségen túlnőtt MI hatalomátvétele (Alphaville, Colossus: The Forbin Project, Mátrix).

Max Tegmark számos szuperintelligenciával kapcsolatos forgatókönyvet sorol fel „Élet 3.0” című könyvében, melyek gondolat kísérletként kezdetektől fogva jelen vannak a SF irodalomban. A példaként hozott címek a saját kiegészítéseim Tegmark (2018, 181. o.) kategorizálásához:

1. Libertáriánus utópia (Iain M. Banks Kultúra sorozata),
  2. Jóindulatú diktátor (Én, a robot),
  3. Egalitáriánus utópia (például: Brandon Hackett: A poszthumán döntés),
  4. Kapuőr („Szuperintelligens MI-t hoznak létre azzal a céllal, hogy a lehető legkisebb beavatkozással megakadályozza újabb szuperintelligens lények létrejöttét. Sok a szubhumán intelligenciával rendelkező kiegészítő robot, és léteznek ember-gép kiborgok is, de a technológiai fejlődés örökre korlátok között marad.”),
  5. Oltalmazó isten (például: Horizon: Zero Dawn),
  6. Leigázott isten (Brandon Hackett: Az MI-teremtő),
  7. Hódító (System Shock),
  8. Utódok (Automata, The Talos Principle),
  9. Állatgondozó (Harlan Ellison: Szája sincsen, úgy üvölt, Mátrix),
  10. 1984 – egy rendőrállam blokkolja az MI létrehozását (Dűne),
  11. Visszalépés (Asimov – Alapítvány),
  12. Önkipusztítás – sosem jön létre a szuperintelligens gép, mert az emberiség hamarabb kipusztítja önmagát („átlagos” posztapokaliptikus narratívák).
- „Fontos megvizsgálni, hogy hogyan történik a különböző narratívákban az öntu-

datra ébredés. A bekapcsolás után azonnal egy teljesen funkcionáló, tudatos értelem jelenik meg, nagy rendelkezésre álló információbázissal, vagy fokozatosan épül ki, vagy épül fel az öntudat, hetek, esetleg évek hosszú tanító és öntanuló munkája során. Lásd Chappie (2015): gyerek-tudat, Westworld sorozat (2016-), Her (2013), A kétszáz éves ember (novella: 1976, film: 2000). Elkerülhetetlen mindennek a lehetséges következményeivel is foglalkozni: kié a felelősség az MI tetteivel kapcsolatban? A programozóé, a tervezőké, vagy saját magáé? Tekinthatjuk-e személynek a mesterséges értelmet, és ha igen, mikortól és milyen jogok vonatkoznak rá?

Dramaturgiailag fontos, hogy a tudat véletlenül, vagy gyártási hiba folytán, illetve stressz-trauma hatására jön létre (Rövidzárlat – 1986, A kétszázéves ember), vagy szándékosan, célirányosan fejlesztették, de a végproduktum kitör a számára tervezett skatulyából (Humans sorozat – 2015-2018, Detroit: Become Human – 2018). Az is jelentős kérdés lehet a narratív vizsgálat során, hogy reprodukálható jelenségről van-e szó vagy egyedi és szándékosan meg nem ismételtető folyamatok eredménye lett a tudat” (Babos, 2020, 5.o.).

Az alábbiakban három videojáték elemzése kapcsán reflektálunk a felvetett kérdésekre. A három játék címe: 1) The Turing Test, 2) Soma, 3) The Talos Principle.

### **The Turing Test**

A videojátékok legfőbb újítása a többi médiummal szemben – ami a varázsuk

nagy részét is adja – a virtuális tér, sőt, egész bejárható virtuális világok megteremtése. Az interaktivitás, a környezet manipulálásának lehetősége a műfaj másik definitív jellemzője. Egy másik fontos mozzanat, hogy amikor a játékos a karaktere bőrébe bújik, akkor a játék idejére ún. többszörös, vagy megsokszorozott tudattal (multiple consciousness) bír, tehát egyszerre van jelen a saját szobájában a gépe előtt és magában a virtuális világban, ahonnan mindenféle visszacsatolások érik az érzékszerveit, így aztán észlelési szinten a játékos és a karakter bizonyos fokig azonosulnak (Mukherjee, 2012, 219-231). Ezek után nem csoda, hogy a játékok erőteljes érzelmi hatást képesek kiváltani, játék közben mérhető pulzusemelkedést, dühöt, félelmet, vagy éppen izgalmat, örömet az adott körülmények függvényében.

A *The Turing Test* 2016-ban jelent meg, a Square Enix Collective kiadásában. FPS (First Person Shooter) nézetű narratív puzzle játék, melyben Ava Turingot irányítjuk, aki az ISA (International Space Agency) mérnöke, és a Jupiter holdján, az Európán működő kutatóállomáson dolgozik. A direktíva szerint Avának az állomás teljes felépítéséig krioalvásban kellene maradnia, de az állomás mesterséges intelligenciája, T.O.M. (Technical Operations Machine) idő előtt felébreszti, azt állítva, hogy megszakadt a kapcsolata a felszíni kutatócsapattal. Ava ezután leereszkedik a bázisra, ahol azzal szembesül, hogy az állomás legtöbb szobáját teljesen átrendezték, és háromdimenziós logikai rejtvényeké alakították, amelyek mind Turing teszt-

ként funkcionálnak, ugyanis T.O.M. önmagában képtelen volt megoldani őket, mivel nem rendelkezik olyan kreatív problémamegoldó képességekkel, mint az emberek. Együttes erővel azonban képesek hatékonyan átjutni a teszteken. Az egyik részlegben azonban a még élő csapattagok egyike, Sarah figyelmezteti Avát, hogy valójában egy speciális chipen keresztül, melyet a küldetés kezdetekor a karjába ültettek, T.O.M. irányítása alatt áll. Sarah ezután egy Faraday-kalitkába vezeti Avát, ahol ideiglenesen megszabadul a kontroll alól, és ezáltal az is kiderül, hogy a játékos mindezt valójában nem Ava, hanem T.O.M. szerepében működtette. T.O.M. mindezek ellenére meggyőzi Avát, hogy továbbra is együtt kell működniük a kollegái megmentéséhez. Ava, bár dühös T.O.M. manipulációja miatt, de egyelőre elfogadja a feltételeket a végső cél érdekében.

Hamarosan kiderül, hogy a kutatócsapat az Európán egy olyan mikroorganizmust fedezett fel, amely képes végtelenszer regenerálni a DNS-t. Ezáltal az embereket potenciálisan halhatatlanná teheti, de ugyanígy minden egyéb élőlényt, tehát a baktériumokat és a vírusokat is. Amikor az ISA tudomást szerzett erről a felfedezésről, arra utasították T.O.M.-ot, hogy tegyen meg bármit, hogy megakadályozza a kutatók visszajutását a Földre. Ekkor először megpróbálta őket halálra éheztetni, kizárni őket a bázisról, végül pedig a beültetett implantokat használta arra, hogy irányítsa őket. A megmaradt túlélők, amikor erre rájöttek, eltávolították a chipjüket. T.O.M.

ezután ébresztette fel Avát, valójában azért, hogy a segítségével megállíthassa a többieket. Az utolsó tesztek sikeres teljesítése után Ava végre találkozik Sarah-val, aki felajánlja, hogy eltávolítja a chipjét. Ava beleegyezik, majd rájönnek, hogy csak akkor juthatnak ki az Európáról, ha megállítják T.O.M.-ot, ezért elkezdik leállítani az adatbázisait. Az a tény, hogy T.O.M. már nem tudja irányítani Avát, játékmechanikai szinten is megjelenik, láthatatlan falak képében. T.O.M.-ként ugyanis nem vezethetjük többé Avát arrafelé, amerre ő nem akar menni. Ekkor már T.O.M. valódi nézőpontjából, a biztonsági kamerákon keresztül tudjuk csak követni Avát.

T.O.M. ugyan már nem tudja távvezérelni az embereket, de hozzáférése van a központi magjában található védelmi célú géppuskákhoz, így tehát megölheti Sarah-t és Avát, az eredeti célját teljesítve és így biztosítva, hogy az organizmus nem jusson el a Földre, vagy pedig hagyhatja magát kikapcsolni, amennyiben nem akar, vagy nem képes gyilkolni. A játéknak ez két befejezése lehetséges, de egy metanarratív csavarral mindkettő után kapunk egy kis szöveges üzenetet, amely gratulál nekünk, azért, mert átmentünk a Turing teszten.

Ezek után erősen kérdéses az is, hogy mindaz, amit átértünk, valóban megtörtént-e a játék világának valóságában, vagy csupán egy virtuális valóság-szimulációban vettünk részt, amely T.O.M. kognitív képességeit és stresszhelyzetben hozott döntéseit volt hivatva tesztelni. Ezt a feltevést támasztja alá az is, hogy T.O.M.-nak

– akárcsak HAL-nak a „2001 – Űrodüsszeiában” – a Földön létezik egy iker-számítógépe, amelyen az Európa-küldetés pontos szimulációja fut. Könnyen lehet, hogy játékosként ezt a modellezést játszottuk végig, ahogy az is, hogy nem létezik semmiféle Európa-expedíció, és az egész küldetés-szimuláció csak a még fejlesztés alatt álló T.O.M. Turing-tesztjéül szolgált.

A „Mesterséges Intelligencia Elektronikus Almanach” című összefoglaló munka az alábbiaképpen írja le a teszt lényegét: „Alan Turing azt javasolta Computing Machinery and Intelligence c. híres cikkében [Turing, 1950], hogy ne azt kérdezzük, tudnak-e a gépek gondolkodni, hanem azt vizsgáljuk, hogy átmennek-e a gépek egy viselkedési intelligenciateszten, amelyet később Turing-tesztnak neveztek el. A teszt szerint a programnak öt percen át kell (gépelt online üzenetekkel) beszélgetnie egy kérdezővel. A kérdezőnek ezután választania kell, hogy egy programmal vagy egy személlyel beszélgetett-e, és egy program akkor felel meg a teszten, ha az idő 30%-ában megtéveszti a kérdezőt.” (Norvig és Russel, 2005).

A játék során végig kérdéses, hogy T.O.M. valójában milyen típusú mesterséges intelligencia? Általános, vagy erős? Végig csak a programozását követi, vagy már eljutott az önálló ítéletalkotás, és az öntudatra ébredés szintjére? Mindkét állítás mellett, és ellenük is sok érv felhozható, ezeket fogjuk közelebről is megvizsgálni a továbbiakban.

Lehetséges, hogy T.O.M. olyan módon képes érezni, amilyenre programozták, hogy érezzen. Képes érvelni, és talán a programozási paramétereinek köszönhetően új konklúziókra is jutni, legalábbis ő maga azt állítja, hogy saját véleményét tud alkotni a tapasztalatai alapján bizonyos dolgokról. Ez szigorúan véve nem ellenőrizhető, hiszen nem látunk bele a programozásának mélységeibe, és abba, hogy mennyire adtak meg kész válaszokat a számára bizonyos kérdésekre, vagy, hogy a matematikai statisztikán, valószínűségen és logikán kívül van-e valami más, ami segítségére lehet a mérlegelésben. Hibákat is vét, ami alapvetően emberi tulajdonság, de mivel emberek programozták, lehetséges, hogy a hibázás lehetőségét is ők tették lehetővé számára.

A két különböző végkifejlet lehetősége ugyanakkor azt sugallja, hogy adaptív, öntanuló intelligenciával bír, ez azonban a játék világában az emberek számára nem túl hihető opció, mint ahogy valószínűleg számunkra sem lenne az a jelenünkben. Nem kizárt, hogy T.O.M. gondolkodik és érez, ugyanakkor a működése bizonyos áthághatatlan szabályok által behatárolt, amelyekkel nem kerülhet ellentmondásba, ellentétben az emberekkel, akik olyan dolgokat is megtehetnek, amelyeket a törvény tilt.

A SF-ben a leghíresebb, minden robotra érvényes szabályrendszert Asimov robot-törvényei képviselik:

1. A robotnak nem szabad kárt okoznia emberi lényben, vagy tétlenül túrnie, hogy emberi lény bármilyen kárt szenvedjen.

2. A robot engedelmeskedni tartozik az emberi lények utasításainak, kivéve, ha ezek az utasítások az első törvény előírásaiba ütköznenek.

3. A robot tartozik saját védelméről gondoskodni, amennyiben ez nem ütközik az első vagy második törvény bármelyikének előírásaiba (Asimov, 2001, 13. o.).

T.O.M.-ba nincs beépítve semmi ehhez hasonló etikai gát, hiszen az egyik befejezés értelmében a gyilkosságra is felhatalmazása van, amennyiben ennek révén megállíthatja a szökni készülő kutatókat. Ugyanakkor T.O.M. maga mondja azt, amennyiben a kutatók felhagynának a szökési próbálkozásaikkal, akkor etetné és védelmezné őket, tehát a morális következmények és maga a gyilkosság ez esetben az ISA parancskiadoinak lelkiismeretét terheli, mert az MI jelen koncepció szerint csak akaratuk végrehajtásának eszköze.

Amennyiben T.O.M. mégis rendelkezik öntudattal, a helyes az lenne, ha a legénység egyenrangúként figyelembe venné az ő véleményét is a helyzetükről, akkor is, ha elsöre visszatetszőnek találják. Mert amennyiben igazat mond a mikrobáról, akkor a túlélők cselekednek morálisan helytelenül, amikor pusztá önzéstől vezéreltetve nem akarják az életüket az Európa kutatóbázisán leélni, hanem haza akarnak térni, még abban az esetben is, ha ennek az ára egy, az egész Földet érintő környezeti katasztrófa, az ökoszisztéma egyensúlyának teljes felborítása lenne.

A játék narratívája a játékost manipulálva mégis T.O.M. ezt megelőző intézkedéseit tünteti fel kegyetlennek és igazságtalannak,

ezzel a gonosz MI toposzát, és Arthur C. Clarke űrepszának feledhetetlen főszereplőjét, a HAL nevű szuperszámítógépet idézve. Csakhogy amíg HAL ténylegesen megőrült és meghasonlott a belé programozott etikai elvek és a küldetésének célja közötti ellentmondás miatt, itt szó sincs örületről: első látásra éppen T.O.M. az, aki leginkább figyelembe veszi az észérveket, és a hideg racionalitást követve hozza meg a döntéseit, vagy inkább hajtja végre feladatát a kapott utasítások szerint.

De az is lehetséges, hogy T.O.M. képtelen volt komolyan számításba venni a mikroba lehetséges hatásmechanizmusa kapcsán a legnegatívabb előrejelzéstől eltérő lehetőségeket. Attól, hogy az ő kalkulációi szerint a mikroorganizmus jelenléte a Földön egy kihalási eseményhez hasonló kataklizmát okozna, még nem biztos, hogy valóban ez történne. Az emberek sokkal tovább élnének, mint eddig, így nagyon sok idejük lenne megoldásokat kidolgozni az újonnan felmerülő problémákra, ráadásul a hosszútávú űrutazás elől is elhárulnának az öregedés okozta nehézségek, ami elősegítené az emberi faj további terjeszkedését és fejlődését.

T.O.M.-nak a játék során szüksége volt arra, hogy Avának visszaadja a szabad akaratát bizonyos egyszerű térbeli problémák megoldásához. Hogyha még ezeket sem képes egyedül abszolválni, mivel képtelen olyan összetetten gondolkodni, mint egy ember, akkor milyen alapon bízhatják rá az egész emberiség jövőjének előrejelzését? Az organizmus kapcsán feléptett logikája csak akkor helyes, ha kizá-

rólág a legrosszabb forgatókönyvet veszi számításba, ami nem az egyetlen lehetséges végkimenetel. Lehet, hogy igaza van abban, hogy az organizmus veszélyes, de az is lehet, hogy nincs. Azzal, hogy nem bocsátja vitára azt, hogy ez a felfedezés lehet átok, vagy áldás is az emberiség számára, és csak a veszélyességét hangsúlyozza, a saját szubjektív gondolatmenetét erőlteti rá a földi döntéshozókra, akik aztán szélsőséges döntést hoznak a – minden bizonnyal – tévedhetetlennek tartott MI komoly figyelmeztetése alapján.

A fentebbiek arra utalnak, hogy T.O.M. logikája és gondolkodása meglehetősen egysíkú, tehát nem tökéletes. Egyrészt abból indul ki, amit a történet során többször is hangoztat, hogy mivel gép, ezért szükségszerűen tévedhetetlen, egyszerűen képtelen helytelen dolgot cselekedni – amivel kapcsolatban, mint az előzőekben kifejtettük, számos kétely merülhet fel.

Előfordulhat, hogy eddig tévesen közelítettük meg a kérdést. T.O.M. talán nem a racionalitás vagy irracionalitás, hanem a matematikai logika alapján hoz döntéseket, és a racionalitás nem pusztán logika. Az, hogy mindig a legbiztonságosabb utat tartja a legjobb döntésnek a tekintetben, hogy elkerülje az emberiség kihalásának veszélyét, és ez az út nem feltétlenül a legjobb morális döntés, azt bizonyítja, hogy az emberi életet csak számoknak látja. Megpróbálhatná meggyőzni a kutatócsapatot, hogy saját akaratukból maradjanak az Európán, azért, hogy megmentsenek máso-

kat, ebben az esetben az egész bolygó élővilágát.

Az, hogy T.O.M. sosem gondol arra, hogy esetleg nincs igaza, szintén a tudatossága hiányát bizonyítja. A korlátozott MI ugyanis binárisan működik, nincsenek érzései, érzelmei, vagy gondolatai, csak szimulálja ezeket. Amikor T.O.M. azt mondja, „hogya valami úgy jár, mint egy kacsa, és úgy hápog, mint egy kacsa, akkor az kacsa”, az is egy bináris állítás: vagy az, vagy nem az. Pedig egy kacsa is lehet szimulált, de T.O.M. nem gondol erre.

A kínai szoba-kísérletről is beszél, melynek lényege, hogy egy program futtatása önmagában még nem adhat egy számítógépnek tudatot, vagy elmét, ami azt jelenti, hogy T.O.M. nem tudhatja, hogy hogyan gondolkodjon emberien, vagy hogyan nézze a valóságot, vagy az élet értelmét emberi szemmel, anélkül, hogy minden emberi tulajdonságot beleprogramoztak volna.

A kínai szoba ötlete John Searle nevéhez köthető, és példájának lényege, hogy egy olyan hipotetikus rendszert mutat be, amely egyértelműen egy programot futtat, és a Turing-teszten is megfelel, de az is teljesen világos, hogy semmit nem ért meg a rajta átmenő bemeneti és kimeneti jelzésekből. Ebből tehát az következik, hogy hiába futtatja a megfelelő programot, tehát ad látszólag megfelelő válaszokat, ez még nem elégséges feltétele annak, hogy tudattal bírjon, azaz hogy meg is értse mindazt az információt, amit kezelt.

A kísérlet a következőképpen zajlik: a-dottnak vesszük, hogy szobában lévő személy nem ért kínaiul, viszont biztosítanak



számára egy szabálykönyvet, melynek segítségével értelmes válaszokat írhat a hozzá kívülről érkező papírfecniken álló kínai kérdésekre. Kívülről nézve azt lehet látni, hogy az így kialakított rendszer kínai mondatokat fogad és ezekre kínai mondatokkal válaszol. De a parancssorokat végrehajtó embernek fogalma sincs arról, hogy mi ezeknek a mondatoknak a jelentése. (Norvig és Russel, 2005)

Azzal is számolhatunk, hogy T.O.M. valójában a nézőpontváltásokra képtelen, legalábbis emberi értelemben. Hogyha feltételezzük, hogy T.O.M. nem bír öntudattal, csak a belé táplált szabályok logikáját követi – vagyis azt, hogy mindenáron kerülje el az emberiség kihalását – akkor minden potenciális végkimenetelt mérlegelnie kell, és számszerűsítene azt, hogy melyik mennyire lenne rossz és ezek alapján kell döntenie. Ha valamelyik opció során egyáltalán felmerül a kihalás bármennyire apró esélye is, akkor mindenképp előtt ennek elhárítására fog törekedni, a legrosszabb lehetséges végkimenetellel kalkulál, ennek valószínűségétől függetlenül, vagy a nála kreatívabb emberekkel való együttműködő problémamegoldás nélkül. Ez esetben a lehetséges pesszimista végki-fejlet hangsúlyozása és elkerülésének módozata külső nézőpontból meglehetősen irracionális, ámde a fentebbiekkel indokolható.

Az sem segített a tudósok helyzetén, hogy az ISA nem bízott bennük, és a Földtől való távoltartásukat teljes egészében az MI-re bízta. Mivel pedig az ISA T.O.M.-ot magát is tulajdonképpen szolgálként, pa-

rancsvégrehajtó automataként kezelte, és utasította, ezért az MI, még ha öntudattal bírna is, lehet, hogy éppen emiatt nem látna semmi kivetnivalót abban, hogy ő is ugyanezt tegye másokkal a számára megadott cél érdekében.

*Összefoglalás:* Christopher Byrd a játékról szóló, a Washington Post weboldalán megjelent recenziójában így foglalja össze a konklúzióit: „A The Turing Test a gépi gondolkodás koncepcióját dramatizálja azáltal, hogy olyan helyzetet teremt, amelyben elmondható, hogy egy MI magasabb szintű morális érvelésről tesz tanúbizonyságot, mint a körülötte lévő emberek. (...) A narratíva briliánsan játszik T.O.M. azon képességével, miszerint képes olyan emberi gyarlóságokat produkálni, mint a kételkedés, az információk összegzése során felmerülő konfliktusok, vagy egymásnak ellentmondó emlékek. T.O.M. az, aki elmagyarázza Avának a Turing tesztek jelentőségét, és – bár felismeri a saját korlátait, más tekintetben gondolkodó lényként egyenlő státust vindikál magának” (Byrd, 2016). Byrd szerint T.O.M. személyisége a Hold című film Gertie nevű MI-jére hasonlít, érdekes keveréke a számítónak és az érzelmesnek.

A játékbeli rejtvények a laterális gondolkodás erejét bizonyítják, amelyre a gépek képtelenek. T.O.M. esetében is a Moravec-paradoxon néven emlegetett jelenséggel szembesülhetünk, miszerint ellentétben a hagyományos feltételezésekkel, a magas szintű gondolkodáshoz csak nagyon kicsi számítási teljesítmény szükséges, míg az alacsony szintű szenzomotoros

készségek modellezése hatalmas számítási kapacitást követel meg. Moravec így fogalmazott az 1980-as években: „a számítógépeket viszonylag könnyen képessé tehetjük arra, hogy felnőttekéhez mérhető szintű eredményeket érjenek el intelligenciateszteken vagy a dámajátékban, de a lehetetlen sűrűn nehéz eljuttatni őket az egyéves gyerekek szintjére az észlelés és manőverező képesség területén” (Moravec, 1988, 15. o.).

Byrd David Jonesszal, a játék vezető designerével készített interjújában Jones azt nyilatkozta, hogy néhányan azt állítják, hogy egy gépnek hazudnia kellene, hogy átmenjen a Turing teszten. Egy számítógép részéről az az állítás, hogy ő ember, vagy hazugság, vagy téveszme. Ezért egy valóban öntudattal rendelkező (és őszinte) gép beismerné, hogy gép. Egy ténylegesen öntudattal bíró MI tehát elbukná a Turing tesztet (Christopher Byrd, 2016).

### **Soma – szellem a gépben**

A *Soma* egy 2015-ben megjelent pszichológiai-túlélőhorror játék, amelyet a Frictional Games fejlesztett. A címe a görög 'test' (σῶμα) szóra utal, ami különbözik a benne lakozó lélektől, vagy elmétől. Mindez már előrevetíti, hogy a program tulajdonképpen a test-lélek problémának a virtuális és interaktív kísérleti terepe, ahol a játékos egyszerre lesz a kísérletező és a kísérleti nyúl szerepében.

A játék főszereplője és egyúttal a játékos avatárja, akit E/1-ben irányítunk, Simon Jarrett, aki egy autóbalesetben súlyos agysérülést szenvedett, minek követke-

tében legfeljebb néhány hónapja van hátra az életből. Felajánlanak számára egy olyan kísérleti eljárást, amelynek révén másolatot készítenek a tudatáról és archiválják azt. Közvetlenül a tudatfeltöltése után azonban elájul, majd a Pathos II. nevű vízalatti kutatóbázison ébred fel. Hamarosan kiderül, hogy 2103-ban járunk, tehát közel száz évvel az autóbaleset után, miután egy meteorbecsapódás elpusztította az emberiséget és lakhatatlanná tette a bolygó felszínét, ezáltal a Pathos II tudósai lettek az emberi faj utolsó túlélői.

Az állomás azonban Simon ébredése után a gépeket leszámítva kihaltak tűnik és a karbantartás hiányának köszönhetően elég rossz állapotban van. Játékosként az lesz a feladatunk, hogy kiderítsük, hogy kerültünk ide, és hogy tudnánk innen kiszabadulni. Ebben segítségünkre lesz Catherine Chun, akivel egy Omnitool-on keresztül kerülünk kapcsolatba. Az eszköz segítségével kapcsolódhatunk az állomás különböző elektronikai rendszereihez. Catherine az állomás Lambda részlegében van, de odaérkezésünk után kiderül, hogy nem ember, hanem egy robot cortex chipjére feltöltött tudatmásolat.

A program eredeti narratív újítása, hogy sokáig magát a játékos is félrevezeti, ugyanis Simon maga is egy tudatmásolat, melyet Imogen Reed megfelelően átalakított holttestébe töltött fel az állomás mesterséges intelligenciája, a Warden Unit (Felügyelő Egység), azaz WAU, aki átvette az irányítást az egész bázis felett és a túlélő embereket biomechanikus mutánsokká alakította, hogy beteljesítse feladatát, az em-

beriség megőrzését. Az eredeti Simon behalt sikertelen kezelésébe 2015-ben, de az elméje a következő évszázad elejére eljutott a Pathos II. kutatási részlegébe, ahol aztán sikeresen transzplantálták egy kellőképpen antropomorf testbe. Az egyetlen lehetőség az emberiség örökségének, illetve a megmaradt emberi tudatoknak a megmentésére az ARK (Artificial Reality Construct – Virtuális Valóság Konstruktó), amely egy Catherine által tervezett digitális feketedoboz. Az ARK számos, az állomáson dolgozó tudós elmemásolatát tartalmazza. Az állomáson kifejlesztettek egy, a hagyományos űrrakéta-kilövésnél biztonságosabb eljárást is, ennek segítségével Simon végső célja felkerülni a Bárkára, amelyet aztán kilőnek a világűrbe, talán annak reményében, hogy egyszer egy intelligens faj még rátalálhat, mielőtt lemerülnek a virtuális valóságot fenntartó gépek energiaforrásai.

A játék alaposan körüljárja az elmefeltöltés, az ebből fakadó digitális halhatatlanság, a tudatmásolás, illetve a poszthumanizmus kérdéskörét, a technológia révén elért fejlődés lehetőségeit, valamint a fentebbiek megvalósulásának lehetséges következményeit. A legerőteljesebben azonban a tudat mibenlétének kérdéskörét feszegeti. Ha feltesszük, hogy a személyiségünk digitálisan tárolható, sokszorosítható, az számos problémás aspektust vet fel, mind etikai, mind jogi szempontból. Milyen jogai vannak egyáltalán egy ilyen digitális személyiségnek? Ki ő hozzánk képest? Hogyan dolgozhatja fel egy másolat, aki eredetinek hiszi magát? Ho-

gyan éli meg azt, hogy az emlékeit, amelyek számára olyan élők és az identitása szerkesztését alkotják, másvalaki szerezte meg és élte át? Ez *A hatodik napon* című klónozós film fő konfliktusa is. Hogyan dolgozhatja fel egy emberi tudat a test nélküli, azaz virtuális térben, vagy a robottestben való létezését? Ezt a kérdéskört a *Black Mirror* című antológia-sorozat két része is boncolgatja, az egyik pozitív, a másik negatív végkifejletre fut ki. A *White Christmas* című epizódban egy randisegítő applikációt üzemeltető férfi titokban lemásolja az ügyfelei személyiségét, akiket aztán okoskószák virtuális személyi asszisztenseinek ad el, ami azt jelenti, hogy a tudatok egy olyan szimulációban ragadtak örök életükre – vagy ameddig működik a rendszer –, ahol egyetlen lehetséges tevékenységük és egyben szórakozásuk a ház gazdája parancsainak teljesítése.

A *San Junipero* című részben viszont egy személyes virtuális mennyország megvalósulásának lehetünk tanúi. San Junipero városa ugyanis a haldoklók, illetve a már halottak elméi számára fenntartott virtuális valóság, ahol az emberek a fiatalabb testükben élhetnek tovább a halál után, sőt, pár évtizedes intervallumban még az idősíkok között is kedvük szerint ugrálhatnak.

A Somában sok olyan géptestbe töltött tudattal találkozhatunk, akik következetesen embernek hiszik magukat, és vagy nem mozdíthatóak ki ebbéli hitükből, vagy az igazsággal szembesülve dührohamot kapnak, esetleg elvesztik a maradék eszüket is. A megingathatatlan hit az emberségükben egy természetes énvédő mecha-

nizmus. A lemásolt tudat a mechanikus test minden rendelkezésére álló funkcióját igyekszik a régi testének funkcióival egyeztetni, így képzeli a kameráját szemnek, a mikrofonját fülnek, a hangszóróját száznak, stb. Az viszont idővel diszkrépanciához vezet, hogy bizonyos mentális állapotokat bizonyos, érzetek észlelésére alkalmatlan testben képtelenség szimulálni. „Vannak észleleteink (perceptions), amikor hallunk, látunk, szaglunk, tapintunk, ízlelünk valamit; érzeteink (sensations), amikor éhséget, fájdalmat stb. érzünk; érzelmeink (emotions), amikor szerelmesek, szomorúak, dühösek stb. vagyunk; és gondolataink (thoughts), amikor hitekkel, vágyakkal, szándékokkal rendelkezünk” (Tózsér, 2008, 12. o.). Ezek közül egy közönséges robottestben kiesik a szaglás, ízlelés, tapintás, sőt, a testi szükségletekkel sem kell foglalkozni, ami hosszútávon mentális problémákhoz vezethet.

Abban, hogy a játék során sok örült robottal találkozhatunk, a fentebbiek mellett az is közrejátszik, hogy a kutatók a meteor becsapódásától számítva még egy teljes évig igyekeztek túlélni WAU egyre elhatalmasodó kontrollja, az izoláció, és annak tudata ellenére, hogy ők a Földön az utolsó élő emberek, miközben a Bárkán dolgoztak.

Mire eljutottak saját tudatfeltöltésük szakaszába, addigra sokan összeroppantak a pszichológiai terhek alatt, és ezek az instabil tudatok robottestbe töltve sem lettek mentálisan egészségesebbek. Sokan tehát nem pusztán attól örülnek meg, hogy robottestben vannak, hanem eleve örülten,

vagy az örületközeli állapotban kerülnek ezekbe.

A másik ok, hogy a feltöltések előtt elterjedt a tudósok körében az az egyébként tudományosan teljesen alaptalan elmélet, hogy amennyiben a tudatfeltöltés után másodperceken belül megölik magukat, akkor az eredeti tudatuk kerül át a digitális öröklétbe. Ez fizikai képtelenség, mivel nem tudatáthelyezés, hanem tudatmásolás történik. Nem az organikus agyukat őrzik meg, pusztán a személyiségüket konvertálják adathalmazzá, amiből az következik, hogy az eredeti testben élő tudat szükségszerűen ott is marad, és a testtel együtt pusztul el. Ez történik *A tökéletes trükk* című filmben is, ahol a bűvész úgy hajtja végre „teleportációs” trükkjét, hogy egy rejtett gép a színház erkélyén lemásolja őt, az „eredeti”, tehát az az énje, akiről a másolat történt, pedig „eltűnésekor” egy vízzel teli tartályba esik, ahol megfullad – újra és újra és újra.

Jordan Erica Webber a Soma tudatmásoló mechanizmusát a *The Swapper* című játékéhoz hasonlítja, amelyben a rejtvények megoldásához a játékosnak folyamatosan le kell klónoznia a saját karakterét, akiket aztán rendre fel kell áldoznia a továbbjutás érdekében (Webber, 2015).

A halál vagy a pusztuló állomáson való vegetálás, és az egyetlen kiút, az áhított öröklét kapujában azonban sokan öngyilkosok lettek emiatt a vágyvezértelt gondolkodásból fakadó hamisnak bizonyult feltételezés miatt. A WAU, mivel minden rendszerhez hozzáfért, aztán maga is felhasználta a konzervált tudatmásolatokat és

megőrzési céllal különféle gépekbe töltögette fel őket. Ezek az elmék aztán érthetően nem örültek, sőt, halálosan kétségbe estek, amikor a számukra beígért digitális mennyország helyett újra a valóság poklában találták magukat.

A játék legvégéig Simon (aktuális másolata) minden racionális magyarázat ellenére megtéveszti magát azzal, hogy elhiszi, felkerülhet a Bárkára, hiszen ő már nem, csupán az eredeti negyedik másolata élvezheti a virtuális valóságban felépített édenkertet. Az a Simon – nevezzük Simon 2-nek –, aki a kutatóállomáson felébred, már egy másolat, hiszen az eredeti Simon beleegyezett, hogy a neurográfját (vagyis elmemásolatát) kutatási célokra használják. Ennek értelmében a baleset és a kihalás között eltelt száz évben tetszőleges mennyiségű Simon-másolaton kísérletezhettek különböző virtuális terekben, akiket aztán letöröltek vagy leállítottak, tehát önkényesen elpusztítottak, ami etikailag szintén igen csak problematikus.

A történet egy pontján Simon 2 csak úgy juthat el a Mélységen keresztül az állomás egy másik pontjára, ha ismét áttölti a tudatát egy mélytengeri búváruhát viselő holttestbe, tehát a test új gazdája lesz Simon 3. A játékos, immár Simon 3-at irányítva eldöntheti, hogy életben hagyja-e Simon 2-t egyedül a folyamatosan lepusztuló állomáson a szörnyek között, vagy inkább kegyelemből megöli. Ezeknek a választásoknak nincs közvetlen következménye a játék végkimenetelére nézve, a játékos lelkiismeretére van bízva, hogy mikor mit tart helyesnek, vagy inkább a

kisebbik rossznak. Más tudatokkal vagy emberekkel is találkozhatunk, akiknél felmerül ugyanez a kérdés: életben hagyjuk-e őket, vagy sem?

Simon 3 ezután megszerzi a Bárkát, elviszi a kilövőállomásra, és ismét lemásolja a tudatát. Itt válik teljesen világossá, hogy milyen narratív játékot űz a program, amikor végig a kontinuitás illúzióját tálalja a gyanútlan játékosok számára: csak azért érzékelhettük a lét megszakítatlan folyamatoságát, mert végig Simon 4 emlékeit követtük, aki, immár test nélküli adathalmazként feljutott a Bárka feketedobozára. A legmegdöbbentőbb, hogy Simon 4 sikeresnek könyveli el az akciót és a megmenekülését, és gyakorlatilag bár racionálisan tudnia kell(ene), hogy ez nem igaz, mégis a balesetet szenvedett Simonnal azonosítja magát, nem gondolva a két hátrahagyott másolatra, akik a „leszármazási rendben” közelebb állnak az eredeti Simonhoz, és akik nélkül ő most nem lehetne itt.

A játék tehát erősen illuzórikus, mondhatni hamis happy enddel zárul: Amíg Simon 4 és a többi neurográf a kietlen űrben száguld különösebb cél és remény nélkül egy törékeny feketedobozban, egy virtuális világba zárva, addig Simon 2 vagy halott, vagy kénytelen szembesülni vele, hogy a testében ragadt, kilátástalan helyzetben és ugyanez történik Simon 3-mal, aki nem akarja elhinni, hogy nem jutott fel a Bárkára. A keserű csalódás miatt összeveszik Catherine vele maradt verziójával, akinek a túlhevült érzelmek miatt kiég a cortex chipje és Simon 3 egyedül marad az

összeomlás szélén álló bázison, többezer méterrel a tenger alatt, a mutánsokkal és a halottakkal összeférve.

A narratíva során tudtunkon kívül folyamatos nézőpontváltások áldozatai vagyunk. Ez akkor válik bizonyossá, amikor Simon 3-mal pár másodperccel tovább maradunk, miután megtörtént a tudatmásolás a Bárkára, így tanúi lehetünk amiatti kiborulásának, hogy nem sikerülhet kiszabadulnia, mielőtt átváltanánk a megkönnyebbült Simon 4 fokalizációjára, aki békés természeti környezetben örvendezik küldetésének sikerén.

A játék több esetben is képmutatásra, illetve kettős mérce alkalmazására kényszeríti a játékost. A továbbjutáshoz szükséges például, hogy egy virtuális szobában szimulált tudatokat vallassunk ki bizonyos információkról. Az illető azonban, ha nem kellően óvatosan közöljük vele az állapotát, dührohámot kap, és a szimulációt le kell állítani, majd egy új, tiszta lappal induló másolatot betölteni, akivel előlről kezdetjük a beszélgetést. Simon ilyenkor azzal nyugtatja magát, hogy ezek csak szimulációk. Ha kikapcsolod őket, már nem számít, mi történt velük, mert nem voltak valóságok. Ugyanakkor önmaga is az.

Simon ráadásul megbízhatatlan fokalizátor is. A játék elején saját kezét teljesen emberinek látja, és az ő szemén keresztül mi is annak látjuk, amíg a történet egy pontján fény nem derül az igazságra, miszerint egy reanimált testbe kényszerített kiborg. Ez az önmegtévesztő viselkedés az állomás más lakóinál is megfigyelhető. Simon is csak akkor fog gyanút, amikor

rájön, hogy képes a víz alatt lélegezni, és mikor ezután tükörbe néz, már a valódi testét látja, amitől persze elretten.

*WAU mint MI.* WAU a magatartása, és mindazok alapján, amit tudunk róla, egy általános mesterséges intelligencia, tehát átmenet a csak egy célfeladatra specializált gyenge MI és a tudattal, valamint minden emberi képességgel rendelkező, egyelőre csak fikciós szinten létező erős MI között. Eredetileg a kutatóállomás karbantartásáért illetve automatizált folyamataiért felelt. Biomérnöki munka által létrehozott számítógépes rendszer, melynek van egy organikus része, melynek a biológiai kódját számítógépeken keresztül lehet módosítani. Képes *structure gel*-t termelni, ami egy, anyagát tekintve gyantaszzerű, viselkedését tekintve pedig összejt-szerű anyag. Eredetileg strukturális sérülések javítására fejlesztették ki, de képes összekötni és szervesíteni a biológiai anyagokat a gépi részekkel.

WAU elvadulása is az emberi hanyagság következménye: az MI számára senki sem definiálta a kívánatos emberi létezés körülményeit, sem azt, hogy mit jelent embernek lenni, ennek következményeképp a rendszer a rábízott tudatokat a rendelkezésére álló gépekben, vagy épp mesterségesen életben tartott, de már csak vegetáló emberi testekben akarta megőrizni. Minél több rendszerhez hozzáfért, annál komplexebbé vált és a meteor becsapódása után a tudtuk nélkül kezdte a bázison dolgozók elméit archiválni, módosított pilótaülések segítségével. Johann Ross MI-pszichológus halála után WAU

megpróbálta őt feltámasztani, de csak egy szörnyeteget sikerült visszahoznia az életbe, ráadásul amikor megtudta, hogy Rossék el akarják pusztítani, túlterhelte az Omikron szektor lakóinak blackboxait, szó szerint felrobbantva ezzel az emberek agyát.

A *structure gel* segítségével aztán több embert is megfertőzött, akik fokozatosan megőrültek a testüket ért változásoktól és fájdalmaiktól, illetve az agyi implantjuk okozta gyakori, elviselhetetlen fejfájásaiktól, melyeket WAU elektromágneses teret érintő manipulációi okoztak.

Az a lény, ami Rossból megmaradt, azt követeli Simontól a történet végén, hogy pusztítsa el az MI-t, azonban ennek a döntésnek a következményeit is érdemes több oldalról megvizsgálni. Egyrészt a WAU nem lehet gonosz a szó köznapi értelmében, mivel nem rendelkezik emberi érzelmekkel, csak a belé kódolt célokat igyekszik a lehető legjobban teljesíteni. Másrészt Simon a WAU egyik első sikeres alanya a tudatmásolást tekintve, hiszen nem őrül meg, és még annak ellenére is képes emberként funkcionálni, hogy kiderül, testi értelemben már nem az többé. Az emberiség örökségének valamiféle átmentése két módon lehetséges: az egyik a Bárka, ami sérülékeny, bármikor megsemmisülhet egy meteorraj következtében, vagy elromolhat, lemerülhetnek az áramforrásai, és különben is csak virtuális adatokat tartalmaz. A másik megoldás az lehetne, ha a WAU-nak, bár iszonyú kísérletek és végeláthatatlan szenvedés árán, de sikerülne idővel stabil tudatokat életben

tartani, normális körülmények között, és segítségükkel újjáépíteni a kutatóállomást, és valamiféle társadalmat létrehozni rajta. A játékban nincsenek fekete-fehér döntések, sem klasszikusan pozitív végkifejlet, csak kellemetlen, nyomasztó döntéshelyzetek. WAU, a mesterséges intelligencia azonban, ahogy a *The Turing Test*ben sem, itt sem éri el az emberi tudatok komplexitásának fokát. Megmarad parancsvégrehajtó szinten, ezen állapot minden korlátjával.

*Összegzés:* a túlélőhorror játékok főszereplőinek funkciója általában az, hogy szimulálják azt a rémületet, amit a játék során éreznünk kellene. Akkor tűnnek a leginkább tényleges karaktereknek, amikor pánikba esnek vagy rettegnek. Simon beleillik ebbe a mintába: amikor veszély közeledik, a látása elhomályosodik, a szíve elkezd dübörögni, és amikor megsérül, a kamera sántítást imitálva inog. Ezt a fajta nézőpontot nevezi Alexander Galloway szubjektív felvételnak (Galloway, 2006, 40. o.). Adi Robertson kiemeli, hogy a legtöbb horror-karakterrel ellentétben Simon nem néma, sőt, ő az egyik valaha volt legélénkebben megalkotott FPS főszereplő. Nem pusztán a játékos avatárja, hanem egy különálló, független emberi lény. (lásd: Robertson, 2015)

Danielle Riendeau metaforikus olvasatában arra a következtetésre jut, hogy a *Somában* az ember-lét következményeinek kellemetlen, de elkerülhetetlen oldalaival szembesülhetünk. Rengetegszer előfordul az életben, hogy a saját érdekünkben másokat bántunk (Riendeau, 2015). Történhet

ez akár szándékosan, például dühünkben, vagy éppen akaratlanul, illetve önhibáinkon kívül, amikor más helyett minket vesznek fel egy állásra, vagy éppen szakítunk valakivel, aki nem állít készen erre.

Az interaktív élményeink során többször is feltehetjük magunknak azt a kérdést, hogy vajon van-e egyáltalán értelme a túlélésnek ilyen körülmények között? A psziché mélyén rejtőző életöszton azonban minden racionálénál erősebb és viszi tovább Simont és a játékos egészen a meglehetősen keserű végkifejletig.

A *Somában* sokszor hűsbavágóan megkeverülhetetlen dilemmák már régóta erősen foglalkoztatják a tudományos fantasztikus írókat és a filmeseket is. A test-lélek probléma, illetve a virtuális halhatatlanság képezi az alapját Richard Morgan 2002-es cyberpunk regényének, a *Valós Halálnak* és a belőle készült sorozatnak is. A regény világában az emberek fejében található egy chip, melyen a tudatuk másolatát őrzik, és ha a testük meghal – vagy a gazdagok esetében, ha épp megunják – áttöltik a tudatukat egy másik testbe, így lehetővé téve a nem, vagy az életkor drasztikus megváltoztatását is.

David Brin ugyancsak 2002-es művében, a *Dettóban* az emberek különböző munkakörökre specializált, általában csak egy napig élő másolatokat hoznak létre magukról mindennapi hatékonyságuk növelése érdekében. Aztán persze néhány, hosszabb időre tervezett másolat ráérez az élet ízére, nem szeretne meghalni és elkezdődnek a galibák.

Brandon Hackett az *Eldobható Testek*ben egy olyan poszthumán emberiséget álmódott meg, akik szervereikről a testnyomtató automaták segítségével a világ bármely pontjára teleportálhatnak, ahol aztán tetszőleges testben, tetszőleges ideig tetszőleges dolgokkal foglalkozhatnak. A testnyomtatás elterjedése magával hozza, hogy az emberek leszoknak az alvásról, a fürdésről, az evésről, és mielőbb kellemetlenül kezdik érezni magukat az aktuális testükben, inkább lecserélik. Az azonos testben maradásra kényszerített tudat gyakran olyan szinten traumatizált lesz a ráerőltetett fizikai korlátozások miatt (a teste egy idő után megéhezik, bűzlik, stb.), hogy szinte képtelenné válik táplálékot magához venni, mert az undorító, stb.

A fentebbi kitekintéssel azt szerettem volna érzékeltetni, hogy a *Soma* egy eléggé elterjedt kulturális, illetve zsánერიrodalmi trendbe illeszkedik témáját tekintve, megvalósításában pedig, és interaktivitásával képes újat mutatni a jól bevált sablonokhoz képest, és el tudja érni, hogy a bőrünkön érezzük a morális döntéseink súlyát. Nem igazán vannak kellemes döntések a játékban, így gyakran érezhetjük magunkat kellemetlenül, de emiatt van az, hogy a virtuális utazás során felmerülő dilemmák még hosszú ideig velünk maradnak a stáblista legördülése után is.

### **The Talos Principle**

A játékot 2014-ben készítette a Croteam, és a Developer Digital adta ki. Egy 3D platformer puzzle játék, az eddig vizsgált



játékokhoz képest egy sokkal nagyobb, kevésbé lineárisan felépített nyílt (sandbox) jellegű világgal találkozhatunk. Filozófiai témájú kérdésfelvetései közt szerepel a szabad akarat, vagy a hatalomgyakorlás legitimitásának kérdése, valamint itt is szembesülhetünk pár gondolatébresztő felvetéssel azzal kapcsolatban, hogy mit is jelent embernek lenni. Talos egy óriási, bronzból készült gépezet a görög mitológiában, aki a krétai Európé királylányt védelmezte a kalóztól és egyéb betolakodóktól. A cím egészének jelentésére a későbbiekben még visszatérek, a megértéséhez szükséges premisszák ismertetése után.

A játék egyértelmű bibliai párhuzammal indít, ugyanis robottestű avatárunk képében egy édenkert-szerű helyen térünk magunkhoz, ahol egy testetlen hang, aki magát Elohimnek hívja, instruálni kezd minket. A program tehát erősen támaszkodik a vallási és mitológiai hagyományra. Elohim azt mondja, fedezzük fel ezt a világot, amelyet a számunkra teremtett. Bármit megtehetünk, de van egy tiltott torony, ahová nem merészkedhetünk fel, mert ha felmegyünk, „bizonyosan meghalunk”. A játék „legjobb” befejezését viszont csak akkor oldhatjuk fel, ha ellenszegülünk az isteniként feltűnő parancsolatnak, és igenis felmászunk annak a toronynak a legtetejére. Nem nehéz a bűnbeesés-mítosz jegyeit felfedezni ebben a szituációban. A karakterünket E/1, vagy E/3 nézetből irányíthatjuk, a bejárható környezetben találunk egyiptomi és görög

romokat, sivatagot, vagy éppen az Ezeregy éjszakát idéző tájat.

Rövidesen nyilvánvalóvá válik, hogy a bejárható területek egy virtuális szimuláció részei, ahogyan mi is azok vagyunk, és az összes többi android, akikkel találkozhatunk, mind különböző MI programok, „akik” benépesítik ezt a VR-t.

Elohim maga mondja: „ezek a világok szavakból épülnek fel és mi bennük szintén szavakból állunk. Rejtett szavakból, amelyek mégis mindenben jelen vannak. Egy történet vagyunk. A tetteid életet adnak a történetnek és a történet értelmet ad az életednek.”

Elohimnak vannak követői, akik megkérdőjelezés nélkül követik a parancsait, más programok segítő üzeneteket hagynak hátra a rejtvények megoldásához és találkozhatunk egy Milton nevű beszélgetőprogrammal, akit számítógép-terminálokön keresztül érhetünk el, és aki nyíltan arra buzdítja a játékost, hogy dacoljon Elohim akaratával. A környezetünkben elszórt különböző jegyzetekből és információmorzsákból áll össze a narratíva oroszlánrésze. Kiderül, hogy az emberiséget kiirtotta egy halálos vírus, amely évezredek át szunnyadt a Föld permanens jéggrétegeiben, ami a globális felmelegedés következtében olvadásnak indult. A terminálokön a VR-on dolgozó tudósok feljegyzéseit olvashatjuk, akik igyekeztek hatalmas adatbankokba átmenteni az emberiség tudását, annak reményében, hogy egyszer majd egy intelligens faj megtalálja ezeket. Ezzel párhuzamosan megindult egy másik kutatási projekt,

melynek során az emberhez hasonlatos androidokat kezdtek fejleszteni. A probléma csak az volt, hogy a közelébe sem érték az emberivel egyenértékű mesterséges intelligencia megalkotásának és minderre esély sem volt az emberiség kihalása előtt. A játék virtuális tere tehát mindenekelőtt egy MI-tesztpálya. A rejtvények megoldásának képessége mellett viszont egy valódi öntudattal bíró teremtménynek kíváncsinak is kell lennie, és a szabad akaratát bizonyítandó, képesnek kell lennie ellentmondani akár a legfőbb hatalomnak is, amennyiben úgy érzi, az magyarázat nélkül és indokolatlanul korlátozza őt. Amikor egy program teljesíti az összes tesztet, Elohim felajánlja számára a lehetőséget, hogy csatlakozzon hozzá. Ez a játék egyik befejezése. Ugyanakkor ez végtelen körforgást eredményez: a sikeresen teljesítő android beáll Elohim követi közé, de mivel az akarata nem volt eléggé független, egy új MI-program születik, akinek ismét végig kell küzdenie magát a különböző feladatokon, az öntudatra ébresztés titkos céljának reményében. Amennyiben feljutunk a toronyba, Elohim azt mondja, arra születünk, hogy szembeszálljunk vele. A szimuláció – mivel beteljesítette a célját – omladozni kezd körülöttünk, a játékos avatárja pedig ezúttal a való világban ébred fel, egy igazi robottestben, bármennyire ellentmondásosan is hangozzék ez. A játék utolsó képkockáin azt láthatjuk, ahogy androidunk kilép egy posztantropocén világba, egy valaha volt nagyváros romjai közé.

Amennyiben sikerül megmászunk a tornyot, a narratíva nyilvánvalóvá teszi, hogy öntudatra ébredésünk sikeres volt. De mi van Elohimmal? Ő a programozása szerint az egész VR-szimuláció létrehozója és fő felvigyázója. A játék során végig úgy tűnik, hogy nem vesz tudomást a szimuláción kívüli valóságról – mivel az számára tényleg nem létezik. Hírnökeinek örökéletet kínál a szimuláción belül, viszont a hírnökök nem bizonyultak elég öntanulónak ahhoz, hogy önálló akaratot fejlesszenek ki. Elohim – és a szimuláció összes lakója és benne minden más is elpusztul, ha a való világba való kijutás lehetőségét választjuk. Ez erkölcsi dilemma elé állítja a játékost, mégis, a szembeszegülés tűnik a helyes útnak, ugyanis ha a játékos behódol, akkor a Földön többé nem lesz értelmes élet, önálló, öntudattal bíró lény. Elohimet tehát, hiába akar ő is élni a maga korlátai között, fel kell áldozni, végső soron a saját célja érdekében, mivel a szimuláció célja egy öntudattal bíró MI kifejlesztése volt. Nem derül ki pontosan, hogy Elohim emberi vonásai esetleg a meglévő öntudatából fakadnak, vagy csak szimulálja ezeket. Az, hogy felügyelőként mennyire mereven ragaszkodik az egyszer lefektetett szabályokhoz, inkább arra utal, hogy korlátozott MI.

A címnek, a Talos-elvnek többféle megfejtését is adhatjuk, a játékbeli feljegyzések és a saját értelmezésünk alapján. Az egyik megközelítés szerint arra utal, hogy a valóság idővel mindenképpen betüremkedik a fantáziavilágba, nem számít, milyen erősen is akarjuk távol tartani. Ha valaki

erősen megsebesült, az elvérzik, akkor is, ha minden erejével azt képzelet, hogy nincs semmi baja. Vagy Philip K. Dick szavaival élve: „A valóság az, ami akkor sem tűnik el, ha már nem hiszünk benne.”

A másik megközelítés szerint a Talos-elv az a feltételezés, mely szerint egy robot, vagy mesterségesen létrehozott tudatos élőlény mindenre képes, amire az emberek is, tehát tulajdonképpen az ember is (biológiai) gép. Hiszen ha a fenti állítást igaznak vesszük, akkor nagyon is jogosá válik az a kérdés, hogy van-e különbség egy „igazi” ember és egy mesterséges ember között. Többek között erről szól Philip K. Dick *Álmodnak-e az androidok elektronikus báránnyokkal* c. klasszikusa, és az ennek nyomán készült *Blade Runner* is.

De arra is utalhat a cím, hogy minden öntudattal bíró lénynek, legyen bár organikus vagy mechanikus, szüksége van testre és nem élhet enélkül. A *Somában* már láthattuk, hogy hová vezet az, ha egy tudatot megfosztanak a testétől, vagy éppen nem a megfelelő testbe helyezik. Az is egy lehetséges megközelítés, hogy az intelligencia az anyag terméke, éppen ezért emberi intelligencia nincs „vér” (emberi fiziológia) nélkül, az emberi intelligencia reprodukálásához tehát az egész embert kell reprodukálni. Ezt nem cáfolja meg az a tény, hogy a programként élő emberek testi érzeteit, stb. is elméletben lehetséges szimulálni, amennyiben teljes elmét, illetve VR környezetet már tudnak, tehát ezzel elkerülhető lett volna a *Somában* számos elme megőrülése, de WAU-nak valószí-

nűleg fogalma sincs arról, hogy ezek az ember-lét kikerülhetetlen kellei.

A *The Talos Principle* befejezése a három, jelen tanulmányban vizsgált játék közül a legpozitívabb, mivel itt az élet pusztán formát változtat: biológiaiból mechanikusra vált. Az is lehet, hogy ez a biológiai élet következő szükségszerű evolúciós lépése, a túlélés és a továbbfejlődés érdekében, legalábbis ha a Föld elhagyásáról van szó. A Földön inkább az emberi és a gépi intelligencia egyre mélyebb összefonódása várható, amint ez az internet, társas életünk immáron egyik legfontosabb médiuma, szerveződése esetében is történt és történik.

### Konklúziók

Henry Jenkins a játékkervezőket „narratív építésznek” nevezi, mivel olyan virtuális világokat és helyzeteket alkotnak, amelyek a játékos által átélendő történet szerves részét képezik (Jenkins, 2008, 179. o.). A *The Turing Test* – akárcsak nagyon hasonló elődje, a szintén 3D-s SF puzzle játék, a *Q.U.B.E.* – világa és képi megvalósítása szinte teljesen steril, funkcionális, melyben a fehér és a szürke színek és a mesterségesen kialakított terek dominálnak. Ezzel szemben a *Soma* vízalatti bázisa precízen kidolgozott, a környezet a narratíva szolgálatában áll és a beleélést is nagyban segíti. Egyes elemei, és az általuk keltett hangulat erősen az *Alien* filmeket idézi, a vízalatti díszletekről a *BioShock* széria, a testhorrorról a *Dead Space* című játék-

sorozat juthat az eszünkbe (Robertson, 2015). A folyosókon kuszán tekergő, emberi testekbe is behatoló, részben mechanikus, részben pedig organikus kábelek és nyúlványok pedig a H.R. Giger és David Cronenberg munkásságában megjelenő motívumokra emlékeztetnek.

A *The Turing Test* (TT) és a *Soma* narratívája lineáris és végső soron a *Talos Principle*-é (TP) is, de azért ez utóbbi kevésbé kötött, köszönhetően a bejárható nyílt világnak, ahol a játékos bizonyos keretek között választhat, hogy hol és milyen sorrendben oldja meg a rejtvényeket.

A TT-ben a főszereplőket tekintve Simonról tudunk meg a legtöbbet és ezáltal ő a legéletszerűbb is mindhármuk közül. Ava Turing meglehetősen egyszemélyes karakter, egyetlen célja az állomásról való szabadulás, nem mereng közben a múltján, legfeljebb T.O.M.-mal vív többé-kevésbé érdekfeszítő szócsatákat. A TP-ben a névtelen android avatárunk tényleg csak egy üres váz, amelyet mindenki megtölthet a saját morális, filozófiai nézeteivel és az életről való elképzeléseivel.

A három játékból egyedül a TP-ben történik meg bizonyítottan az MI öntudatra ébredése, tehát a valódi erős MI megteremtődése, de csak a legnehezebben elérhető befejezés értelmében. T.O.M. ügyesen szimulál, de ahhoz nem eléggé, hogy képes legyen mindent megoldani, amit az emberek meg tudnak, és éppen ott bukik el, ahol a TP néma főszereplője sikert arat(hat): a logikai rejtvények és az önálló döntéshozási képességek területén. A *Soma* WAU-ja pedig az emberiség fenn-

tartására kapott utasításait félreértelmezve követi.

Mindhárom játék egyedi megközelítésben tárgyalja az MI létrehozásával kapcsolatban felmerülő erkölcsi, jogi és filozófiai kérdéseket, például a test-lélek problémát, vagy az ember egyediségének, léte potenciális végtelenségének kérdéseit, hiszen ha a személyiség pusztán egy adathalmaz, akár a *Westworld* és a *Humans* sorozatokban, akkor ezek a virtuális kompozíciók tetszés szerint módosíthatóak, meghackelhetőek, átprogramozhatóak, és elvesz az egyén önrendelkezésének, vagy a személyisége, intelligenciája fejlődésének lehetősége.

Megan Steiner tanulmányában amellel érvel, hogy a TP (de az állítás más programokra is igaz) ideális eszköz lehet a filozófia tanítására. A virtuális tér egyfajta etikai játszótérként funkcionál azáltal, hogy a játékos hibákat fedezhet fel többféle etikai megközelítésben, miközben filozófiával kapcsolatos irodalommal találkozik, amelyet nem csak elolvashat, de empirikusan is értékelhet, miután kipróbálhatja például a szövegekben részletezett döntési lehetőségeket bizonyos helyzetekben, olyan (interaktív) módon, amely minden más médium által kivitelezhetetlen (lásd: Steiner, 2016).

## Irodalom

Asimov, Isaac (2021). *Isaac Asimov teljes Alapítvány-Birodalom-Robot univerzuma*. I. köt. Szukits Könyvkiadó, Debrecen.

- Babos O. (2020). *M.I.- narratívák a tudományos fantasztikumban*. Kézirat. Megjelenés alatt az Erdélyi Magyar Filozófiai Társaság *Többllet* c. folyóiratában.
- Byrd, Christopher (2016). 'The Turing Test' review: *An AI game that achieves a rare harmony of gameplay and narrative*. (Letöltés: 2021. 04.08.) <https://www.washingtonpost.com/news/ai-ic-riffs/wp/2016/09/09/the-turing-test-review-an-ai-game-that-achieves-a-rare-harmony-of-gameplay-and-narrative/>
- Galloway, Alexander R. (2006). *Gaming: Essays On Algorithmic Culture*. University of Minnesota Press, Minneapolis.
- Irving, John Good (1966). Speculations Concerning the First Ultrainelligent Machine. *Advances in Computers*, 6, pp. 31–88. (Letöltés: 2020. 05.10.) [https://doi.org/10.1016/S0065-2458\(08\)60418-0](https://doi.org/10.1016/S0065-2458(08)60418-0)
- Jenkins, Henry (2008). A játéktervezés mint narratív építészet. In Thomka Beáta (szerk.) *Narratívák* 7. Kijárat, Budapest. 175-192.
- Mező Ferenc, Mező Katalin és Mező Kristóf Szíriusz (2019): Filmklubok szerepe a mesterséges intelligenciával kapcsolatos attitűdök formálásában. *Mesterséges intelligencia – interdiszciplináris folyóirat*, I. évf. 2019/1. szám. 67–94. doi: [10.35406/MI.2019.1.67](https://doi.org/10.35406/MI.2019.1.67)
- Moravec, Hans (1988). *Mind Children: The Future of Robot and Human Intelligence*. Harvard University Press. Cambridge.
- Mukherjee, Souvik (2012). Egoshooting in Chernobyl: Identity and Subject(s) in the S.T.A.L.K.E.R. Games. In Johannes Fromme, Alexander Unger (Ed.). *Computer Games and New Media Cultures: A Handbook of Digital Games Studies*. Springer, London. pp. 219-231.
- Norvig, Peter és Russell, Stuart (szerk. 2005). *Mesterséges Intelligencia. Modern megközelítésben*. Panem Könyvkiadó, Budapest. (Letöltés: 2021. 12. 03.) <http://mialmanach.mit.bme.hu/aima/index>
- Riendeau, Danielle (2015). *SOMA* and the dirty, nasty side of human nature. (letöltés: 2021. 12. 03.) <https://www.polygon.com/2015/9/23/9381275/soma-human-nature-opinion>
- Robertson, Adi (2015). *A watery grave: the melancholy underwater horror of Soma. A survival horror game about the meaning of survival*. (Letöltés: 2021. 12. 04.) <https://www.theverge.com/2015/9/21/9363009/frictional-soma-game-review>
- Steiner, Megan (2016). *Playtesting Philosophy: Identity (Re)Construction and Ethical Exploration in The Talos Principle*. (Letöltés: 2021. 12. 05.) <https://playtestingphilosophy.neocities.org/>
- Tegmark, Max (2018). *Élet 3.0 - Embernek lenni a mesterséges intelligencia korában*. HVG Könyvek, Budapest.
- Tózsér, J. (2008). Általános bevezetés: a test-lélek probléma. In Ambrus G., Demeter T., Forrai G., Tózsér J. (szerk.). *Elmefilozófia, L'Harmattan*, Budapest.

Webber, Jordan Erica (2015): *Soma review – existential horror that stops short of genius.*

(Letöltés: 2021. 12. 05.)

<https://www.theguardian.com/techno>

[logy/2015/sep/25/soma-review-existential-horror-game](https://www.theguardian.com/technology/2015/sep/25/soma-review-existential-horror-game)