

„Szia, Alexa!” A hangvezérlés mint új technológia kulturális megalkotása a platform-ökoszisztémában

TÓFALVY TAMÁS*

Az utóbbi évtizedben egyre nagyobb teret nyer a hangvezérlés, más néven hangalapú interfész a digitális-média-ökoszisztémában. A számos platform által bevezetett legkülönbözőbb szoftver-hardver konfigurációk közül a legismertebb az Amazon Alexa, a Google Home és az Apple Siri. A nagy platformok a hangalapú interfészeket jellemzően „virtuális asszisztensként” vezetnek és mutatják be, kommunikálják a felhasználók felé. A hangvezérléssel működő virtuális asszisztensek a számítógépen és az okostelefonon kívül más digitális eszközökkel is összekapcsolhatók, kiterjesztve ezáltal a hangvezérlés hatókörét a hétköznapi élet számos területére. A tanulmány egyfelől azt elemzi, hogyan zajlik a technológia kulturális megalkotása, azaz a technológiáról szóló közbeszédben milyen értékek, jelentések és figuratív nyelvi alakzatok kapcsolódnak a hangvezérléshez (például „okos” vagy „asszisztens”). Másfelől azt vizsgálja meg, hogyan illeszkedik ez a fajta jelentésalkotás és antropomorfizáció a platformok által dominált digitális-média-ökoszisztéma működésébe, hogyan támogatja a hangalapú interfész a minél szélesebb körű felhasználói adatgyűjtést és -monetizációt célzó platformlogikát.

Kulcsszavak: digitális média, hangalapú interfész, hangvezérlés, kritikai technológiakutatás, platformok, virtuális asszisztensek, VUI

„Hi Alexa!” The cultural construction of the voice user interface as a new technology in the platform ecosystem

Voice control, or Voice User Interfaces, have been gaining ground in the digital media ecosystem over the last decade. Among the wide variety of software-hardware configurations introduced by many market players, some of the best known are Amazon Alexa, Google Home, and Apple Siri. The major platforms typically brand and introduce voice interfaces as ‘virtual assistants’ to the market and the users. Voice-controlled virtual assistants can be connected to a variety of digital devices other than computers and smartphones, extending the reach of voice control to many areas of everyday life. In this paper, I critically analyze how this technology is culturally constructed, i.e. what values, meanings and linguistic tropes (e.g. metaphors such as ‘smart’ or ‘assistant’) are associated with Voice User Interfaces in public discourse about the technology. On the other hand, I will explore how this kind of meaning-making fits into the functioning of the platform-dominated digital media ecosystem, how Voice User Interfaces support platform logic aimed at collecting and monetizing as much user data as possible.

Keywords: digital media, voice interface, voice control, critical technology studies, platforms, virtual assistants, VUI

* Egyetemi docens, Budapesti Műszaki és Gazdaságtudományi Egyetem.

1. Bevezetés

2014. november 6-án új videó került fel az Amazon YouTube-csatornájára.¹ A kisfilmen az volt látható, hogy egy családhoz megérkezik egy titokzatos fekete henger, amelyhez beszélni lehet, és az – női hangon – válaszol is. Mindenféle kérdést fel lehet neki tenni („Hány óra van?”, „Milyen magas a Mount Everest?”), és a legkülönbözőbb utasításokat is rögtön végrehajtja („Játssz rockzenét!” vagy „Adj csomagolópapírt a bevásárlólistához!”). A készüléknek szánt kommunikációt egy megszólítással („Alexa”) kell jelezni, hogy az egyébként mindig bekapcsolt állapotban lévő, a felhővel és a kiválasztott eszközökkel online összeköttetésben lévő eszköz az utasítások teljesítését és a kérdések megválaszolását az interneten keresztül végrehajtsa.

A videó az aznap piacra dobott Amazon Echo nevű „okoshangszóró” reklámfilmje volt, a készülék lelke, azaz a beszédfelismerésre és -alkotásra is alkalmas „virtuális asszisztens” szoftverre pedig az Alexa nevet kapta. A hangalapú interfészként (*voice user interface*, VUI) működő Alexa ma már nemcsak az Echo-ban, hanem több ezer további készülékben is megtalálható, és az Amazon sincs egyedül a piacon. Az egymással versenyző VUI-k közé tartozik például az egyik legrégebbi szereplőként az Apple Siri, a Google Home, valamint a 2021-ben leállított Microsoft Cortana is. A különböző platformok által alkalmazott hangalapú interfészek vagy éppen hangvezérlés által működtetett virtuális asszisztensek megtalálhatók játékkonzolokban, okostelefonokban, tévékben, karórákban, összekapcsolódva megannyi további eszközzel, akár a hűtőszekrénytől egy automatizált lakás világító- vagy fűtőrendszeréig.²

Az egy évtizede az Amazon Echo által újraélesztett és azóta folyamatosan növekvő és terjeszkedő hangalapú interfész piacának jelentőségét egyfelől az adja, hogy az a felhasználói számítástechnika 1980-as évekbeli, valamint a hétköznapi internethasználat 1990-es évekbeli kezdeteitől fogva az ember-számítógép interfészben (*human-computer interface*) domináns vizuális/grafikus (*graphical user interface*, GUI),³ majd a 2010-es évektől kezdve az érintéssel kiegészített grafikus interfész mellé zárkózik fel újabb interfészmodozatként. Másfelől az, hogy mindez a mostanában a digitális világ oligopol uraivá váló platformok, techóriások felügyelete alatt történik, a platformizáció részeként, összekapcsolódva a mesterséges intelligenciának nevezett megoldások legújabb hullámával.⁴

¹ Introducing Amazon Echo. Know Your Meme, <https://bit.ly/3OrUXqQ>.

² Áttekintéshez lásd Carmen BĂLAN: Chatbots and Voice Assistants: Digital Transformers of the Company–Customer Interface. A Systematic Review of the Business Research Literature. 18(2) *Journal of Theoretical and Applied Electronic Commerce Research* (2023), <https://doi.org/10.3390/jtaer18020051>, 995–1019.; Bettina MINDER et al.: Voice Assistants in Private Households: A Conceptual Framework for Future Research in an Interdisciplinary Field. 10 *Humanities and Social Sciences Communications* 173 (2023), <https://doi.org/10.1057/s41599-023-01615-z>; Francis RAKOTOMALALA et al.: Voice User Interface: Literature Review, Challenges and Future Directions. 1(2) *System Theory, Control and Computer Journal* (2021), <https://doi.org/10.52846/stcej.2021.1.2.26>, 65–89.; Simone NATALE – Henry COOKE: Browsing with Alexa: Interrogating the Impact of Voice Assistants as Web Interfaces. 43(6) *Media, Culture & Society* (2021), <https://doi.org/10.1177/0163443720983295>, 1000–1016.

³ Lev MANOVICH: *The Language of New Media*. Cambridge, MIT, 2001, <https://doi.org/10.22230/cjc.2002v27n1a1280>.

⁴ Felix M. SIMON: Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy. 10(10) *Digital Journalism* (2022), <https://doi.org/10.1080/21670811.2022.2063150>, 1832–1854.

Ebben a tanulmányban azt vizsgálom, hogyan illeszkedik a hangvezérlés a platformok által dominált digitálismédia-ökoszisztéma működésébe. A folyamat jelentőségének megértéséhez elsősorban a technológiai fejlesztések és a piaci működés társadalmiasulásán és kulturális megalkotásán keresztül közelítetek. A kortárs technológiakutatás konstruktivista felfogása adja a vezérfonalat, amely arra fókuszál, hogyan működnek közre az egyes (főként az „újnak” nevezett) technológiák észlelésében, elterjedésében, befogadásában, használatában a kulturális konstrukciós mechanizmusok, például a jelentések, az értékek⁵ és a nyelvi eszközök. Ezért a fő kérdés egyfelől az, hogy a technológiáról való közbeszédben milyen értékek, jelentések és figuratív nyelvi kifejezésmódok (például metaforák) kapcsolódnak a hangvezérléshez, másfelől pedig az, hogy ez a folyamat hogyan függ össze a platformizációval: hogyan valósítja meg és támogatja a platformlogikát a hangalapú interfész térnyerése? Elsőként egy rövid történeti áttekintésben mutatom be a mai hangvezérlés piac fő működési elveinek közvetett és közvetlen előzményeit, valamint a platform-ökoszisztémába való beágyazottságát, majd az ebben a kontextusban megfigyelhető kulturális megalkotás módozatait vizsgálom.

2. Hang: média, interakció, interfész

A hang és az elektronikus média és kommunikáció bár több mint egy évszázados közös múltra tekinthet vissza, a történet korántsem töretlen ívű. Évtizedekig a telefonos hangalapú kommunikáció volt a meghatározó a személyközi kommunikációban, és a hangrögzítés és hangfelvételek sugárzása által jött létre a mai értelemben vett zeneipar is. Azonban a mai értelemben vett, a hétköznapok részeként működő digitális, majd később online média történetének kezdeteikor (tehát nem idesorolva például a zenestúdiókban már az 1970-es években elterjedt digitális technológiákat) korántsem volt ennyire jelentős a hang szerepe. A személyi számítógépek az 1980-as évek közepétől kezdve bár alkalmasak voltak zenei tartalmak lejátszására, és változatos – mára már ikonikussá vált – hangjelzéseket is kiadtak a számítógép éppen aktuális állapotainak, műveleteinek visszajelzésére, az alapvető tartalomtípus, amelyet a készülék közvetített, vizuális (szöveges és képi) volt. Hasonlóképpen az 1990-es évek közepétől elterjedő online világban az interneten keresztül hozzáférhető tartalmak is elsősorban szövegek és képek voltak, és csak az ezredfordulót (és a szélessáv elterjedését) követően váltak megszokottá a hangalapú és a hangos mozgóképes tartalmak és kommunikációs módok, mint például az internetalapú hangkommunikáció vagy később a videohívások. (Mindemellett külön fejezetet kaphatnának a mé-

⁵ A konstruktivista, avagy kulturális relativista technológiakutatás egyik meghatározó műve Trevor PINCH – Wiebe E. BIJKER: Tények és termékek társadalmi konstrukciója, avagy hogyan segítheti egymást a tudomány szociológiája és a technika szociológiája (ford. KIRÁLY Gábor). *Replika*, 51–52. (2005) 57–87. A társadalmi konstrukció által létrehozott technológiai jelentések vizsgálatát megalapozó mű Carolyn MARVIN: When Old Technologies Were New Thinking About Electric Communication in the Late Nineteenth Century. Oxford, Oxford University Press, 1990, <https://doi.org/10.1093/oso/9780195063417.001.0001>. A konstruktivista/relativista megközelítést történeti perspektívából foglalja össze Gabriele BALBI – Paolo MAGAUDDA: *A History of Digital Media*. London, Routledge, 2018, <https://doi.org/10.4324/9781315209630>.

diatörténetben a hangjelzések is, a legváltozatosabb sípolások, pittyegések, kattogások, amelyekkel az egyes készülékek – illetve azok alkotói – próbálnak üzeni a felhasználónak.⁶)

Ennél is lényegesebb, hogy a tartalomfogyasztás és az ember-ember közötti kommunikáció mellett a felhasználók és a készülékek kommunikációja, az ember-gép interakció az egész médiatörténetet tekintve – az elektronikus média kezdeteitől egészen az ezredforduló online médiájáig – szinte kizárólag a vizualitás és az érintés kombinációjával valósult meg.⁷ Ahogy a GUI segítségével történő tájékozódás, úgy a korai rádiók tekerőgombjának elfordítása a jelölt pozíciók valamelyikébe, vagy a mobiltelefonon egy társkereső alkalmazásban a jobbra vagy balra húzott profil is a grafikus jelek közötti tájékozódás és az érintés általi manipuláció együtműködését valósítja meg.

Ezen a folyamatoságon belül természetesen számos olyan fázis, szakasz elkülöníthető, amelyben a felhasználók tömegei radikálisan újak érezhettek egy-egy változást, és az egy-egy technológiai megoldás elterjedéséhez volt köthető. Ilyen volt például a távirányító, amely lehetővé tette a tévékészülék és a néző fizikai eltávolodását, hozzájárulva ezáltal a tévézési szokások és a családon belüli hatalmi dinamikák alakulásához is (a sok órán keresztül tartó, folyamatos ide-oda kapcsolgatás közben nem mindegy, kinél van a távirányító, azaz a kontroll a tévét nézők között).⁸ De ilyen volt az első iPhone megjelenését követően meghatározóvá váló érintőképernyő is (hiszen korábban is elérhető volt számos így működtethető eszköz), amelynek „újdonsága” abban rejlett, hogy a vizualitással mindig is együtt járó érintés nem a mozgó alkatrészekkel (gombok, egér, billentyűzet), hanem a virtuális gombokat, csúszkákat, billentyűket, kapcsolókat megjelenítő érintőképernyő felületével lépett kapcsolatba. Az adott időszakban akár gyökeresnek érezhető változások az eszközök és azok használati módjai terén – az interfészek típusai felől szemlélve – valójában nem jelentettek alapvető váltást, mert mindegyik trend a vizualitás és az érintésalapú interfészek nagy, évszázados trendjébe illeszkedett.

A mindennapokban használható hangalapú interfész aránylag kései megjelenésének – ahogy a nagyközönség számára nagyon újak ható összes technológiának – szintén évtizedekre visszanyúló előzményei vannak. Akárcsak a jelenleg éppen legújabb felfutását élő, a hangalapú interfésszel összefonódva „virtuális asszisztensként” is megjelenő mesterséges intelligencia legalább hetvenéves történetét, a hangvezérlés technológiájának fejlesztési ívét is számos „nyár” (olyan periódus, amikor a technológia divatos, közérdeklődést vált ki, ezzel összefüggésben megnő a fejlesztésére fordítható és pályázható, ezért sokan foglalkoznak vele) és „tél” tagolta.

A mesterséges intelligenciához hasonlóan az ember-gép kommunikációt kiterjesztő hangalapú interfészek története is legalább az 1950-es évekig vezethető vissza. A Bell Laboratories által 1952-ben tervezett első beszédfelismerő rendszer az Audrey nevet kapta: az Automatic Digit Recognition Machine, azaz automatikus egységfelismerő gép a nevéhez híven diszkrét, elkülönített hangsorokat ismert fel, a számokat pedig nulla és kilenc között, ha a beszélő kellő hosszúságú szünetet tartott közöttük. Az 1970-es években a US Department of Defense Advanced Research Project Agency (DARPA) indította el a Speech Understand Research (SUR)

⁶ Miért kell mindennek sípolnia? *Telex*, 2024. január 5., <https://bit.ly/3AUSPEV>.

⁷ David PARISI: *Archaeologies of Touch: Interfacing with Haptics from Electricity to Computing*. Minneapolis, University of Minnesota Press, 2018, <https://doi.org/10.5749/j.ctt20mvgvz>.

⁸ Ellen SEITER et al. (szerk.): *Remote Control: Television, Audiences, and Cultural Power*. London, Routledge, 1989.

programját (SUR), amely a Carnegie Mellon egyetemmel együttműködésben célozta meg a beszédfelismerés kutatását és fejlesztését. Ennek az együttműködésnek az egyik eredménye lett a Harpy rendszer, amely több mint ezer kimondott szót tudott felismerni (angolul).⁹ Párhuzamosan a számítástechnika új területeinek előtérbe kerülésével, a beszédfelismerés, a természetes nyelvfeldolgozás, a beszédgenerálás és -alkotás kutatásaival és fejlesztéseivel (különféle területeken, például az akadálymentesítésben, az autó- vagy a játékiparban¹⁰) az 1990-es évekre egyre több olyan szolgáltatás, eszköz vált elérhetővé a hétköznapi felhasználók számára, amely vagy diktálásra, vagy a számítógép utasítására, vagy éppen más modalitások hangra fordítására volt alkalmas (például a DragonDictate).¹¹

3. Interfész: vezérlés, asszisztencia, hangszórók

Egy évtizede kezdődött el a hangalapú interfész legutóbbi, még jelenleg is tartó nyara és egyben terjeszkedése a mindennapok digitális ökoszisztémájában. A folyamatnak egyszerre volt elindítója és indikátora is az Amazon Echo 2014-es piacra dobása, ahol már jelen volt az a sajátos fogalmi összefonódás, amely az ilyen interfészek vállalati bemutatását, márkázását, ezzel összefüggésben a médiában megjelenő keretezését, továbbá a közbeszédben való megalkotását jellemzi. Ennek a legszembeütőbb jele az, hogy amikor hangalapú interfészekről beszélünk, akkor nem hangalapú interfészeket emlegetünk, hanem leggyakrabban vagy „virtuális asszisztenseket”, vagy „okoshangszórókat”. Bár a hangvezérlés nem azonos sem a virtuális asszisztensekkel, sem az okoshangszórókkal, a vállalati zsargon, a média és a közbeszéd sokszor szinonimaként használja ezeket a kifejezéseket.¹²

Pedig a három között jelentős különbség van, és mindegyik a digitális média-ökoszisztéma különböző területeivel köti össze a hangalapú interfészeket. A „virtuális asszisztenseknek” egyáltalán nem szükségszerű tulajdonságuk, hogy hanggal kommunikáljanak: virtuális asszisztensnek tekinthető akár egy csetalkalmazás is, amelynek utasításokat lehet adni, az pedig feladatokat végez el vagy opciókat kínál. De éppígy asszisztensnek tekinthető egy olyan navigációs alkalmazás is, amely megtervezi és megmutatja a két megadott pont közötti lehetséges út-

⁹ Joshua ALSPECTOR – Thomas G. DIETTERICH: DARPA's Role in Machine Learning. 41(2) *AI Magazine* (2020), <https://doi.org/10.1609/aimag.v41i2.5298>, 36–48.

¹⁰ Az iparági fejlesztésekről lásd Kwan MIN LEE et al.: Voice Orientation of Conversational Interfaces in Vehicles. 43(3) *Behaviour & Information Technology* (2023), <https://doi.org/10.1080/0144929X.2023.2166870>, 433–444.; Chaiwooo LEE et al.: User Acceptance of Voice Interfaces in the Automobile. 59(1) *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2015), <https://doi.org/10.1177/1541931215591355>, 1641–1645.; Omer TSIMHONI – Daniel SMIT – Paul GREEN: Address Entry While Driving: Speech Recognition Versus a Touch-Screen Keyboard. 46(4) *Human Factors* (2004), <https://doi.org/10.1518/hfes.46.4.600.56813>, 600–610.; Fraser ALLISON – Marcus CARTER – Martin GIBBS: Word Play: A History of Voice Interaction in Digital Games. 15(2) *Games and Culture* (2020) 91–113., <https://doi.org/10.1177/1555412017746305>.

¹¹ George PHILIP – F. J. SMITH – Danny CROOKES: Voice Input/Output Interface for Online Searching: Some Design and Human Factor Considerations. 14(2) *Journal of Information Science* (1988), <https://doi.org/10.1177/016555158801400204>, 93–98.

¹² Smit DESAI – Michael TWIDALE: Metaphors in Voice User Interfaces: A Slippery Fish. 30(6) *ACM Transactions on Computer-Human Interaction* (2023), <https://doi.org/10.1145/3609326>, 1–37.

vonásokat, és a felhasználói preferenciák mérlegelésével kiválasztja ezek közül a legmegfelelőbbnek tartott utat. Virtuális asszisztensek ilyenformán már léteztek a hangalapú interfészek térnyerése előtt is, és továbbra is léteznek azoktól függetlenül is. Egy másik, hasonlóan kurrens *boomját* élő iparág, a generatív mesterséges intelligencia területéről hozva példát, ilyen asszisztensnek tekinthető a felhasználók számára hozzáférhető legtöbb mesterségesintelligencia-alkalmazás a ChatGPT-től egészen a Google Bardig.

Az „okoshangszóró” szintén azért nem lehet szinonimája a hangalapú interfésznek, mert az annak mindössze egy hardveres, ráadásul opcionális része. A hangalapú interfész szoftverek ma már számtalan különböző, hangérzékelésre és -kibocsátásra alkalmas tárgyba lehetnek beépítve, így nem azonosíthatók csak az egyik eszközzel a sok közül. Már csak azért sem, mert nem is kell feltétlenül egy külön hangszóróba telepítve működniük, használhatók anélkül is, például egy okostelefon, egy táblagép vagy egy laptop hardveres közvetítése által.

Mégis, az utóbbi évtizedben a hangalapú interfészek piacát vezető platformok következetesen asszisztensnek és/vagy okoshangszórónak nevezik hangalapú megoldásaikat. Így tett a VUI-fejlesztésekbe a legkorábban belekezdők egyikeként a Microsoft is, amely több alternatív változat (például a Louise Personal Assistant) mellett 2009-ben kezdte el fejleszteni a Cortanát. A 2014-ben be-, majd 2021-ben végleg kivezetett¹³ szoftver mindvégig asszisztensként szerepelt a vállalat kommunikációjában. Az, hogy nem egy okoshangszóróval azonosították, feltehetően azzal a körülménnyel is magyarázható, hogy mindössze egy céggel sikerült megegyezniük arról, hogy saját okoshangszóróikat a Cortanával együtt dobják piacra. A Harman Kardon Invoke Speakerének eladásai ráadásul olyannyira lemaradtak a többi piaci szereplőétől, hogy az árának megfelelését követően nemcsak a készülék gyártását szüntették be, de még a már eladottakat is olyan frissítéssel látták el, hogy eltávolítható legyen róluk a Microsoft szoftvere.¹⁴ Piaci információk szerint a Cortana üzleti felhasználóknak célozva marad a Microsoft-portfólió része, és a felhasználók számára utódját, a Windows Copilotot fejlesztik tovább.¹⁵

A másik korai fejlesztő, az Apple asszisztense eleinte külön dedikált okoshangszóró nélkül vált hozzáférhetővé a gyártó egyes eszközein Siri néven. A 2011-ben bevezetett kommunikációs eszköz számára az Apple hosszú ideig kereste a hardveres eszközt, amely 2015-től a vállalat által megvásárolt Beats Pill lett (2022-ben vezették ki), majd 2018-tól a HomePod vált az első számú dedikált eszközzé, amelyben külön operációs rendszer, az audioOS biztosítja az alapokat.

Ezzel párhuzamosan a Google is több szoftveres és hardveres megoldást fejlesztett a területen, amelyek karrierje változatos módokon metszette egymást. Az egyik első kezdeményezés a vállalatnál a Google Voice Search volt (2012-től), majd a Google Now 2012 és 2014 között, amelynek helyét a szoftveres vonalon a Google Assistant vette át 2016-tól. A hardverek terén az első modell a Google Home okoshangszóró volt 2016-ban, ezt követően fokozatosan a Google

¹³ Microsoft's Cortana Silenced as Siri Gets New Voice. *BBC News*, 2021. április 1., <https://bit.ly/3Z8FTDd>.

¹⁴ Cortana Explained: Rocky Past, Promising Future for Microsoft's Digital Assistant. *Computerworld*, 2022. március 2., <https://bit.ly/494ke3X>.

¹⁵ Microsoft Announces Windows Copilot, an AI 'Personal Assistant' for Windows 11. *The Verge*, 2023. május 23., <https://bit.ly/4fHyHFu>.

Nest vált nem pusztán a központi hardverre, hanem a vállalat által kínált hangalapú megoldások ernyőmárkájává is mint az automatizált háztartások központi egysége és irányítója.

Azonban a legújabb VUI-terjeszkedés piacilag legsikeresebb és legmeghatározóbb szereplője kétségkívül az Amazon, amelynek a 2014-ben bevezetett Amazon Echo okoshangszórója – benne az Alexa szoftverrel – vált a mintává a többi versenytárs számára is. Kevés hivatalos és hiteles adat áll rendelkezésre arról, pontosan mennyi Amazon Echo kelt el eddig és hogy az eladásai hogyan viszonyulnak a szegmens adataihoz. A vállalat 2023-as közlése szerint több mint 500 millió Alexa-szoftverrel ellátott készüléket adtak már el, de azt nem pontosították, hogy ebből mennyit tesznek ki az Echo készülékek.¹⁶ Egyes kimutatások szerint jelenleg 400 millióhoz közeli háztartásban vannak „okosotthon” eszközök, amelyeknek része lehet az „okoshangszóró”. Az Egyesült Államokban (ahol legalább 95 millió ilyen eszköz található) és az Egyesült Királyságban (ahol a lakosság 40 százaléka birtokol ilyen eszközt) a legelterjedtebb a technológia, és a piacot az Amazon vezeti, nyomában a Google-lel. A globális piacon további releváns szereplők az Apple és a kínai gyártók, például a Baidu, az Alibaba és a Xiaomi.¹⁷

A hangalapú interfész, a virtuális asszisztens és az okoshangszóró körüli fogalmi zavarok és átfedések nem esetleges velejárói, hanem szükségszerű következményei a platform-ökoszisztéma működése mögötti logikának. Az elnevezések is részei annak az összetett kulturális keretrendszernek, amelyben bizonyos értékeket, üzeneteket, minőségeket társítunk az egyes technológiákhoz, amelyhez hozzátartozik a termékeknek a fogyasztók felé való vállalati kommunikálása, az értékesítés folyamata, valamint a technológiák társadalmi megalkotásának is meghatározó részét képezik.

4. Kulturális megalkotás: metaforák és antropomorfizáció

A hangalapú interfész technológiájának kulturális megalkotása több, egymással összekapcsolódó szinten zajlik: többek között a nyelv, az esztétikai ítéletek, a morális állásfoglalások vagy az olyan fogalmi műveletek, feltételezések útján, amelyek a technológia természetéről és a világgal való viszonyáról szólnak. Ebbe a körbe tartozik például a technológia kapcsolata az emberrel, a társadalommal, a múlttal, a jövővel, az idővel és a változással.

A nyelvi szinten többek között azok a figuratív nyelvi alakzatok, például metaforák a meghatározók, amelyeket a technológiáról szóló diskurzusokban használunk. Ahogy az már a számítástechnika kezdete óta jelen van a számítógépekről, az internetről és a grafikus interfészekről szóló beszédben (a „mappáktól” a „böngészésig”),¹⁸ a hangalapú interfészekről szóló beszédet is számos metafora alakítja.¹⁹ Ilyen például az „okos-” előtag: ezt szinte minden olyan eszköz vagy infrastruktúra nevéhez hozzátoldjuk, amely valamilyen összeköttetésben áll az

¹⁶ Amazon has Sold More Than 500 Million Alexa-Enabled Devices, Drops 4 New Echo Products. *Yahoo! Finance*, 2023. május 17., <https://bit.ly/495WWKM>.

¹⁷ Smart Speakers – Statistics & Facts. *Statista*, 2024. január 14., <https://bit.ly/416ydnW>.

¹⁸ MANOVICH i. m. (3. lj.).

¹⁹ DESAI-TWIDALE i. m. (12. lj.).

interfészsel: okosotthon, okostelefon vagy okoshűtő. A metafora kognitív megalkotása²⁰ során a forrástartományt (intelligens, gondolkodó entitás tulajdonsága) vonatkoztatjuk a céltartományra (az eszköz adatszolgáltatása, manipulációja): „az eszköz úgy szolgáltat adatot és úgy működik, mintha okos lenne.” A művelet végén pedig már helyettesítjük a céltartományt a forrástartománnyal: „az eszköz okos.” Ugyanez a művelet zajlik le, amikor virtuális asszisztensekről beszélünk: „az eszköz úgy csinál, mint egy asszisztens” asszociációból az „eszköz egy asszisztens” megfeleltetésig jutunk el. Ezek a legalapvetőbb nyelvi szinten, reflexszerűen és tudattalanul, felhasználók milliói által használt alakzatok is beleilleszkednek abba az átfogóbb trendbe, amelyben a platformizációt kiterjesztő technológiai vállalatok arra törekednek, hogy termékeik, szolgáltatásaik a legjobb fényben tűnjenek fel, és azok minél szélesebb körű és gyakoribb használatára motiválják a felhasználókat. A hangvezérléshez kapcsolódó metaforák mellett hasonló szerepet tölt be a platformizáció többi figuratív nyelvi alakzata, metaforája, amelyek célja a céltartományok tényleges funkciójának, működésének elfedése,²¹ az arra irányuló fogyasztói figyelem elterelése a szélsőségesen pozitív vagy éppen semleges forrástartománnyal való párosítás által. Erre lehet példa a „felhő” metafora a „valaki más számítógépe” helyett, vagy éppen a „megosztás” a bérlés vagy kölcsönzés helyett.

A hanginterfészek és kapcsolódó eszközeik nyelvi keretezése, metaforizációja, *brandelése* mellett a techvállalatok más módokon is töreksenek a hanginterfészek önálló, „gondolkodó” ágensekként, emberi létezőkként való bemutatására. A platformok stratégiája itt található a társadalmi képzelettel, az érdeklődéssel és az azzal szervesen összefonódva működő média és szórakoztatóipar felhajtóerejével.²² Az emberi természet antropomorfizációs hajlamának köszönhető, hogy nemcsak a természeti, fizikai vagy biológiai jelenségeknek, hanem a technológiai alkotásoknak, számítási folyamatoknak is emberi jelleget és minőséget tulajdonítunk,²³ ezért a hanginterfészek nagyobb gyártói is már a kezdetektől a legkülönfélébb antropomorfizációs módokkal kísérik meg emberibbé tenni az interfészeket (túl a fentebb említett nyelvi keretezésen).

Ennek egyik viszonylag közvetlen és egyszerű módja a termékek emberként vagy ember-szerű lényként való ábrázolása. Ennek egyik legkorábbi példája az Apple Beats Pill-reklámjai – főleg a debütálása 2013-ban az MTV Video Music Awardson²⁴ –, amelyekben a hangszórók mint emberszerű figurák beszélgettek egymással, az egyik legfrissebb pedig az Amazon Astro háztartási robot, amelynek a képernyőn megjelenített, animált (pislogó, mozgó) szeme és a

²⁰ Zoltán KÖVECSES: *Metaphor: A Practical Introduction*. Oxford, Oxford University Press, 2010.

²¹ Lásd The Misleading Power of Internet Metaphors. *Scientific American*, 2018. szeptember 5., <https://bit.ly/4eRUCZ8>; How Metaphors Shape the Digital Society. *Humboldt Institut für Internet und Gesellschaft*, 2017–2018, <https://bit.ly/3V7TucZ>; Sally WYATT: Metaphors in Critical Internet and Digital Media Studies. 23(2) *New Media & Society* (2021), <https://doi.org/10.1177/1461444820929324>, 406–416.

²² Justine HUMPHRY – Chris CHESHER: Preparing for Smart Voice Assistants: Cultural Histories and Media Innovations. 23(7) *New Media & Society* (2021), <https://doi.org/10.1177/1461444820923679>, 1971–1988.

²³ Simone NATALE: The ELIZA Effect: Joseph Weizenbaum and the Emergence of Chatbots. In Simone NATALE: *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford, Oxford University Press, 2021, <https://doi.org/10.1093/oso/9780190080365.001.0001>, 50–67.

²⁴ Beats Pills VMAs Commercial, <https://bit.ly/3Zr2wo5>.

gesztusai próbálnak emberszerű benyomást kelteni.²⁵ Ennél összetettebb antropomorfizáló stratégiák is megjelennek a termékek hirdeteiben.²⁶ Ilyen például a hanginterfész családi interakciós, kommunikációs hálózatba beemelése, annak részévé tétele. Az Amazon Echo már említett első reklámjában a hanginterfész részt vett a családi szabadidő eltöltésében, bevonódva a társalgásba. Még nagyobb szerepet kapnak a készülékek a Google Nest²⁷ és az Amazon Astro reklámjaiban: nem csupán a családtagokkal emberszerűen kommunikáló résztvevőként, hanem az egymástól távol lévő vagy csak éppen a szomszéd helyiségben lévő családtagokat összekötő szereplőként, a család kommunikációs hálózatának lényegi részeként jelennek meg.

Az interfészek hangdizájnya és a velük végrehajtható interakciók is támogatják a termékek antropomorfizációját. A hang, amelyen az interfész az audio információkat továbbítja a felhasználó számára, emberi hang. Lehetne olyan megoldás is, amely nem emlékeztet emberi hangra, de a fejlesztői szándék terméke gyakorlatilag az összes VUI esetében emberi hang, leggyakrabban női,²⁸ és „mainstream” (tehát nem helyi dialektusban megszólaló) angol nyelven beszélő hang,²⁹ amelyet valódi emberek hangjának felhasználásával állítanak elő. Nemcsak emberi hangjuk van az interfészeknek, hanem emberi módon is beszélnek, ami szintén nem szükségszerű: közölhetnék a kért információk lényegi tartalmát is, ehelyett természetes nyelvi fordulatokban beszélnek. Ez pedig befolyásolja azt is, hogy a felhasználók hogyan értelmezik, fogadják be az információkat és hogyan fogalmazzák meg a VUI felé a kérdéseiket vagy igényeiket.

Amíg egy grafikus interfészen csak kattintunk egyet vagy beírunk egy keresőkifejezést, addig a kimondott utasításokban hajlamosabbak vagyunk informális nyelvi fordulatokba ágyazni az igényeinket, túlmutatva az egyszerű utasításokon. (Ez a beszélgetésformátum írásban is létrejöhét, mint például a chatbotok esetében.) A pusztá parancsvégrehajtáson túli, beszélgetésszerű interakcióként megtervezett ember-gép kommunikációt támogatják azok a köszönési, megszólítási aktusok is, amelyekkel aktiválni kell az (egyébként mindig bekapcsolt állapotban lévő, csendben figyelő) VUI-kat. Ellentétben a grafikus interfésszel, itt „ébresztőszavakkal” vagy aktiváló kifejezésekkel (*wake up word*, *activation phrase*) meg kell szólítani a gépet, hogy használni lehessen. Ez lehet egyszerűen csak a neve kimondása: „Alexa”, vagy egy „Hey, Google” (esetleg „OK, Google”), vagy „Hey, Cortana” beköszönés is.³⁰ Ágensként lépünk interakcióba az interfésszel, miközben nem pusztán a gépet antropomorfizáljuk, hanem áttételesen a mögötte álló márkát is megszemélyesítjük.

²⁵ Introducing Amazon Astro – Household Robot for Home Monitoring, with Alexa, <https://bit.ly/3V8Eonw>.

²⁶ Miriam LIND – Sascha DICKEL: Speaking, but Having No Voice: Negotiating Agency in Advertisements for Intelligent Personal Assistants. 30(3) *Convergence* (2023), <https://doi.org/10.1177/13548565231192100>, 1008–1024.

²⁷ Family Time | Google Home Mini, <https://bit.ly/3Zn2AF3>.

²⁸ Quynh N. NGUYEN – Anh TA – Victor PRYBUTOK: An Integrated Model of Voice-User Interface Continuance Intention: The Gender Effect. 35(15) *International Journal of Human-Computer Interaction* (2019), <https://doi.org/10.1080/10447318.2018.1525023>, 1362–1377.; Rochelle EDWARDS – Philip KORTUM: He Says, She Says: Does Voice Affect Usability? 56(1) *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2012), <https://doi.org/10.1177/1071181312561295>, 1486–1490.

²⁹ Selina J. SUTTON et al.: Voice as a Design Material: Sociophonetic Inspired Design Strategies in Human-Computer Interaction. In *CHI Conference on Human Factors in Computing Systems Proceedings*, 2019, <https://doi.org/10.1145/3290605.3300833>, 1–14.

³⁰ Voice activation, <https://bit.ly/414tAuz>.

A hangvezérlés kulturális keretezésében a nyelvi és az antropomorfizáló képzettársítások mellett a „futurisztikum” és a „forradalmiság” eszméi is szerepet kapnak. Az előbbi szerint a jelenlegi technológia valamilyen módon a jövőt előlegezi meg, felhasználóját a jövőbe helyezi. Ezt az egyébként már logikai szinten teljesen inkonzisztens üzenetet a közelmúltban a leg-hangsúlyosabban a metaverzum bemutatásakor használta a Meta,³¹ de általában a digitális technológiákkal kapcsolatban az egyik leggyakrabban felmerülő kulturális koncepció. A hangvezérlés, a beszéd általi kommunikáció a gépekkel régóta visszatérő motívuma vagy a jövőbe helyezett tudományos-fantasztikus műveknek, vagy ugyan az aktuális jelenben játszódó, de a kor technológiáit „meghaladó” megoldást bemutató alkotásoknak, a *Blade Runner* (1982) replikánsaitól és a *Knight Rider* (1982) KITT-jétől az *Iron Man* (2008) J.A.R.V.I.S.-én át a *Nő* (2013) Samanthájáig.

Az egyes eszközökhöz kapcsolt „forradalmiság” eszméje azt az elképzelést közvetíti, hogy a technológiai fejlődés egy-egy újítás előzmények nélküli megjelenése (esetleg egy „zseni” általi feltalálása) révén zajlik, és a „forradalmian új” eszközök formálják a világot. Ez a felfogás a hangvezérléssel kapcsolatban is igen elterjedt, és a populáris kultúrán kívül a Simone Natale és munkatársai kifejezésével találóan „vállalati determinizmusnak” elkeresztelt ideológiában,³² a techvállalatok kommunikációs holdudvarában tevékenykedő evangélisták, guruk és megmondóemberek (*pundit*) megnyilatkozásaiban, a termékek marketingjében, valamint a médiában is uralkodó szerepe van.³³ Hogyan működnek közre mindezek a nyelvi és kulturális jelentésképzések a technológiák társadalmi megalkotásában és a platform-ökoszisztéma működtetésében?

5. A hangvezérlés a platform-ökoszisztémában

A jelenleg legsikeresebb VUI-k mindegyike valamelyik nagy platform, azaz techcég terméke. Ha nem is mindegyik interfész, alkalmazás vagy készülék a saját fejlesztésük eredetileg, de felvásárlásokon keresztül az Amazon és a Google dominálja a piacot, és a további fejlesztések is globálisan vagy regionálisan meghatározó platformokhoz tartoznak az Apple-től kezdve a kínai Tencenten át egészen az orosz Yandexig. Bár nehéz általánosan érvényes definíciót adni a platformokra, az a technikai vonatkozásuk közös, hogy ezek a vállalatok egy digitális infrastruktúra segítségével a keresletet és a kínálatot kötik össze, „a szolgáltatást generáló vállalkozásokat a fogyasztókkal/felhasználókkal”.³⁴

Ezek a jellemzően oligopol pozícióban lévő vállalatok a platformizáció folyamata során úgy formálják a digitális ökoszisztémát, hogy az egyes fejlesztési trendek, szoftverek és eszkö-

³¹ TÓFALVY Tamás: Meta-metaverzum: Egy technológia diszkurzív megalkotása, a platformizáció logikája és a jövő kérdőjelei. *Szabad Piac*, 2023/1., 53–62.

³² Simone NATALE – Paolo BORY – Gabriele BALBI: The Rise of Corporational Determinism: Digital Media Corporations and Narratives of Media Change. 36(4) *Critical Studies in Media Communication* (2019), <https://doi.org/10.1080/15295036.2019.1632469>, 323–338.

³³ Például James VLAHOS: *Talk To Me: How Voice Computing Will Transform the Way We Live, Work, and Think*. New York, Harper, 2019.

³⁴ GÁLIK Mihály – CSORDÁS Tamás (szerk.): *A média gazdaságtanának kézikönyve*. Budapest, Média tudományi Intézet, 2020, 48.

zők összekapcsolásával a lehető legjobban támogassák és valósítsák meg a platformlogikát, azaz a felhasználói viselkedés és tranzakciók adatainak lehető legpontosabb követését és dokumentálását, és annak alapján a célzott tartalmak, termékek, reklámok közvetítését és értékesítését. Ez a platform- vagy adatkapitalizmusnak is nevezett üzleti stratégia szorosan összefonódik a „súrlódásmentes kapitalizmus” eszméjével.³⁵ A Bill Gatesnél már 1995-ben megjelent,³⁶ majd a platformizációval egyre jobban kiteljesedő koncepció szerint a lehető leghatékonyabb adatgyűjtés, valamint a folyamatos tartalomszolgáltatás és -fogyasztás érdekében az automatizálásnak arra kell törekednie, hogy a felhasználói élmény a lehető legerőfeszítésmentesebb legyen, és hogy a folyamatokat megszakító emberi beavatkozás minél inkább kiküszöbölhetővé váljon. Emellett az egyes platformok úgy alakítják a saját szerepüket ebben a rendszerben, hogy az egyes felhasználó az általuk kialakított alókoszisztémán belül maradjon a digitális élete során a levelezéstől kezdve a térképhasználaton át egészen a böngészésig. A platformlogika által meghatározott mindkét stratégiának alapvető része a hangvezérlés elterjesztése és támogatása. Azt, hogy a hangvezérlés hogyan szövi át egyre több területen és módon a platform-ökoszisztémát, több más tényező mellett alapvetően három párhuzamosan alakuló és változó terület, trend összefonódása mutatja be.

Az egyik ilyen trend – amelynek részét képezik az „okosotthon”-fejlesztések – az internetnek a fizikai világgal való egyre szorosabb kapcsolata. A legkülönbözőbb készülékeknek az interneten keresztül egymással és egy központi feldolgozó és irányító központtal való összekapcsolódását nevezzük a „dolgok internetének” (*internet of things*, IoT). A dolgok internete jelen van az iparban (szenzorok hálózatával a külső körülményeket figyelembe vevő öntözőrendszerektől kezdve az egymással összeköttetésben lévő gyártósorokig), a szolgáltatások területén (például a valós idejű helyzetmeghatározás alapján frissülő menetrendi információkban) és a hétköznapi életben is, legjellemzőbben az otthoni, háztartási eszközök automatizálásában és/vagy távolsági irányításában (az alkalmazásból indítható forró levegős sütőktől kezdve a médialejátszó rendszerekig). A dolgok internete a készülékek felől érkező (egyre növekvő) feldolgozható adatmennyiséget jelenti, valamint a készülékekkel való online kommunikációt és azok irányíthatóságát. A dolgok internetének ezeket a jellemzőit azonosítja a köznyelv az utóbbi években az „okos” jelzővel – ami fontos részét képezi a termékek marketingjének és társadalmiasulásának.

A második meghatározó trend az algoritmizáció és a mesterséges intelligencia, valamint a robotok szerepe részben a háztartások, az ipar és a dolgok internete folyamatainak automatizálásában, részben az audiovizuális és zenei streamingplatformok, valamint a híroldalak forgalmának optimalizálásában és alakításában. A felhasználói viselkedés alapján összeállított Spotify-zeneajánlásoktól kezdve az Amazon Astro háztartásfigyelő kisrobotig az automatizáció egyre inkább lényegi eleme lesz az emberek és a digitális eszközök közötti adatforgalom feldolgozásának, elemzésének és az ezek alapján kialakított vállalati stratégiák végrehajtásának.

A harmadik az okostelefon mint konvergens eszköz központi szerepe az ökoszisztémában. A készülék olyan, mint egy központi interfész, irányítóközpont, amely mindig velünk van, reg-

³⁵ J. Ignacio GALLEGO: The Value of Sound: Datafication of the Sound Industries in the Age of Surveillance and Platform Capitalism. 26(7) *First Monday* (2021), <https://doi.org/10.5210/fm.v26i7.10302>.

³⁶ BILL GATES: *The Road Ahead*. London, Viking, 1995.

gel azzal kezdjük a napot, este pedig magunkkal viszünk az ágyba. Emellett lehetnek még további, speciális feladatokra készített eszközök is, ilyen például az okoshangszóró vagy a Google Nest Hub okoshangszóró-képernyő kombináció, de az okostelefon és a rajta keresztül elérhető többi szolgáltatás és felület szerepe mára megkerülhetetlenné vált.

Hogyan kapcsolódhatnak össze mindezek a folyamatok a hangvezérléssel a platformizáció kiterjesztésében és a platformlogika követésében? A fent kifejtett két fő platformstratégia alapján a nagy oligopol techcégek elemi érdeke, hogy a tevékenységük alapjául szolgáló felhasználói adatok gyűjtését a lehető leghatékonyabban és a legszélesebb körben tudják végezni. Ezért a hangalapú interfész minél nagyobb társadalmi elfogadottsága és elterjedése legalább három szempontból fontos a platformcégeknek. Egyfelől az adatgyűjtés megkönnyítése miatt: a VUI-k alapértelmezetten mindig bekapcsolt állapotban vannak (*always on*), a háttérben hallgatva és gyűjtve az adatokat a felhőbe, azaz távoli szerverekre, ahol azok folyamatos elemzése hozzájárul ahhoz, hogy az adott vállalat hatékonyabbá tudja tenni a szolgáltatásait. A felhasználónak ebben a konstrukcióban gyakorlatilag semmilyen feladatot nem kell végrehajtania, nem kell energiát fektetnie az adat létrejöttébe (mint például egy vásárlás esetén), elég csak élnie az életét a bekapcsolt okos eszköz mellett. Ugyanilyen fontos szempont az adatok mennyiségi növekedése: az audiotérben hatalmas mennyiségű adat halmozódik fel nap mint nap. Az írott kommunikációnk (például az e-mail vagy a chat) régóta a (meta)adatgyűjtés tárgya, de eddig hatalmas mennyiségű beszédat feldolgozatlan maradt. A VUI-k lehetőséget adnak arra, hogy az életünk során keletkező kommunikációs adatok nagyobb része feldolgozhatóvá váljon. Végül a hangalapú interfészek segítségével a platformok olyan újfajta adatokat is gyűjthetnek a mindennapi élet több szegmenséből (az intim- és a magánszférából), amelyekhez eddig nem vagy csak áttételesen volt hozzáférésük. Az adatok újdonsága azok típusára is érvényes, hiszen az érzékszervi kiterjesztés által a szöveges, a képi, az audiovizuális és a biometrikus adatok mellett már a hangunk is a gyűjtött adatok körébe tartozik.

A hangvezérlés alapvető fontosságú a platformlogika másik megnyilvánulási formája, az alókoszisztéma-építés számára is. Ugyanis mindezek a platform-ökoszisztémára általánosan érvényesnek tekinthető stratégiák szervesen összekapcsolódnak az alókoszisztéma-építési törekvésekkel is. Az egyes platformok által kialakított, félig zárt alókoszisztémák lényege az, hogy a felhasználó bejelentkezve – például az Amazon vagy a Google rendszerébe – a lehető legtöbb cselekvést és fogyasztást végre tudja hajtani a rendszeren belül maradván. A felhasználó már hardveres szinten a platform termékéből indulhat ki, például egy Google Pixel okostelefon vagy egy Chromebook megvásárlásával. Ezzel a lépéssel máris elkötelezte magát a Google által kínált operációs rendszer, az Android iránt, majd azt használva szintén további Google-termékek kerülnek előtérbe: ha az interneten keresgélne, akkor azt a Chrome böngészővel (is) teheti, alkalmazásokat a Play Store-on keresztül tölthet le, ha videós vagy zenei tartalmat fogyasztana vagy keresne, ha kommunikálna, navigálna, folyamatokat és napirendet tervezne, netán vásárolna, akkor azokhoz a rendelkezésére áll a YouTube és a Google Music, a Google Search, a Gmail, a Google Maps, a Google Calendar, valamint a Google Wallet. Ha nem a Google, hanem például a Meta alókoszisztémájának a portfólióját vesszük szemügyre, ott a központi felületet a közösségi média (a Facebook és az Instagram) adja, a hozzájuk szervesen kapcsolódó üzenetküldő szolgáltatásokkal (Messenger, WhatsApp). Ugyanez az Amazon esetében az e-kereskedelem köré szerveződik.

Az alökoszisztémák építésének célja az, hogy a felhasználó minél több időt töltsön el a platform felületein és lehetőség szerint ne hagyja el azt. Ennek érdekében a VUI-hoz hasonló érzékszervi kiterjesztést célzott meg a Meta is a metaverzum projektjével, amely – legalábbis a koncepció szerint – az immerzív, mozgásos-audiovizuális élmény kiépítésén keresztül kísérelte meg támogatni a két fő platformstratégiai célt: az adatok minél szélesebb körű gyűjtését és a felhasználónak az alökoszisztémában tartását a lehető legakadálymentesebb módon.

A hangvezérlés technológia – főként az azt előállító és forgalmazó platformok által meghatározott – nyelvi és kulturális megalkotásának egyes szintjei és módjai is ezt a platformlogikát támogatják. Az „okos” jelző kiterjedt használata eltereli a figyelmet a készülékek adatgyűjtő- és továbbító szerepéről, és intelligens, kifinomult, emberi asszociációkat kapcsol a gépekhez. Az antropomorfizációs stratégiák (az emberi hang használata, a beszélgetésformátum alkalmazása, az eszközök családtagként való bemutatása) pedig tovább távolítják egymástól az interfész és a hozzá kapcsolódó eszközök valós célját, az adatgyűjtést és -feldolgozást. Az interfész ágensként kezelése hatékonyan járul hozzá ahhoz, hogy a felhasználók ne kérjék számon szigorúan személyes adataik kiterjedt kezelését a változatos módokon megszemélyesített interfészekben,³⁷ és a gyártótól részben függetlenül kezeljék azokat, egyfajta „digitális belenyugvással”.³⁸

6. Konklúziók

A tanulmányban azt elemeztem, hogy a hangvezérlés mint elterjedő interfész forma működése és jelentései hogyan konstruálódnak meg a mai digitális média-ökoszisztémában. Az ember-számítógép interfész (HCI) domináns formája évtizedeken keresztül a grafikus interfész (GUI) volt, amely mellé az utóbbi években kezd felzárkózni a hangalapú interfész (VUI). A VUI kulturális megalkotása, jelentésadása terén azt mutattam be, hogy bár a VUI-kat jellemzően virtuális asszisztensként vagy tetszőleges okoseszközként kommunikálják a gyártók és forgalmazók, de ezek mindössze az átfogóbb kategória, a VUI, azaz a gép-ember kommunikációt, interakciót hangalapú információ útján vagy annak segítségével megvalósító interfész részei, elemei vagy ahhoz kapcsolódó kiterjesztései. Az interfészekről szóló kommunikáció és a hozzá kapcsolódó diskurzus számos olyan értéket, tulajdonságot tulajdonít a VUI-technológiáknak, nyelvi és fogalmi szinten egyaránt, amely nem része ugyan a technológiáknak, de meghatározza piaci szerepük és társadalmi szerepük jelentőségét is. Például a VUI-k antropomorfizációja során emberi tulajdonságokkal ruházzák fel az interfészt, amely hatással van a velük folytatott kommunikáció jellegére és a technológiának a mindennapi életben betöltött szerepére is.

Ezek a kulturális mechanizmusok szervesen kapcsolódnak a platformizáció folyamatához, részben annak termékei, és így végső soron a nagy platformok célkitűzéseit és érdekeit,

³⁷ Smit DESAI – Michael TWIDALE: Is Alexa Like a Computer? A Search Engine? A Friend? A Silly Child? Yes. *CUI 2022: 4th Conference on Conversational User Interfaces*, Glasgow, 2022. július, <https://doi.org/10.1145/3543829.3544535>, 1–4.

³⁸ Elizabeth FETTEROLF – Ekaterina HERTO: It's Not Her Fault: Trust Through Anthropomorphism among Young Adult Amazon Alexa Users. *Convergence* (2023), <https://doi.org/10.1177/13548565231200337>.

azaz a platformlogikát szolgálják. A platformok célja a szolgáltatásaik és a médiatartalmak szolgáltatása melletti célzott hirdetések értékesítése. Ezzel összefüggésben a platformok számára rendkívül fontos egyfelől a felhasználói viselkedés és tranzakciók adatainak a lehető legpontosabb követése és dokumentálása, valamint az automatizálás által a felhasználói élmény „súrlódásmentessé” tétele. Másfelől pedig az, hogy a felhasználók az általuk kialakított alókoszisztémán belül maradjanak a digitális életük során.

Ennek a platformlogikának és -stratégiának szerves része a hangvezérlés elterjesztése és támogatása, legalább három fő okból. Egyfelől az elengedhetetlenül fontos állandó adatgyűjtés megkönnyítése miatt. A VUI-k készülékei alapértelmezetten mindig bekapcsolt módban vannak és gyűjtik az adatokat a távoli szerverekre. A felhasználónak ez súrlódásmentes adatszolgáltatást tesz lehetővé, nem kell energiát fektetnie az adat létrehozásába. Másfelől a begyűjtött adatok mennyiségi növekedése miatt is. A hangtér beemelése az adatgyűjtés körébe lehetőséget ad arra, hogy a felhasználók élete során keletkező kommunikációs adatok még nagyobb része feldolgozásra kerüljön. Harmadrészt olyan újfajta adatok is gyűjthetők így a mindennapi élet több szegmenséből az intimitás és a magánszféra megfigyelésén és feldolgozásán keresztül, amelyhez eddig nem vagy csak áttételesen volt hozzáférésük a platformoknak. A gép és az ember közötti interfész érzékszervi kiterjesztésével minden eddiginél hatékonyabb és szélesebb körű adatgyűjtés valósítható meg súrlódásmentes élményen keresztül, mindez a technológiákat (és ezáltal a mögöttük álló vállalatokat) barátságos és emberi létezőként bemutató metaforák kulturális-nyelvi konstrukciójába ágyazva.

Irodalomjegyzék

- ALLISON, Fraser – CARTER, Marcus – GIBBS, Martin: Word Play: A History of Voice Interaction in Digital Games. 15(2) *Games and Culture* (2020) 91–113.
<https://doi.org/10.1177/1555412017746305>
- ALSPECTOR, Joshua – DIETTERICH, Thomas G.: DARPA’s Role in Machine Learning. 41(2) *AI Magazine* (2020) 36–48.
<https://doi.org/10.1609/aimag.v41i2.5298>
- BĀLAN, Carmen: Chatbots and Voice Assistants: Digital Transformers of the Company–Customer Interface. A Systematic Review of the Business Research Literature. 18(2) *Journal of Theoretical and Applied Electronic Commerce Research* (2023) 995–1019.
<https://doi.org/10.3390/jtaer18020051>
- BALBI, Gabriele – MAGAUDDA, Paolo: *A History of Digital Media*. London, Routledge, 2018.
<https://doi.org/10.4324/9781315209630>
- DESAI, Smit – TWIDALE, Michael: Is Alexa Like a Computer? A Search Engine? A Friend? A Silly Child? Yes. *CUI 2022: 4th Conference on Conversational User Interfaces*, Glasgow, 2022. július, 1–4.
<https://doi.org/10.1145/3543829.3544535>
- DESAI, Smit – TWIDALE, Michael: Metaphors in Voice User Interfaces: A Slippery Fish. 30(6) *ACM Transactions on Computer–Human Interaction* (2023) 1–37.
<https://doi.org/10.1145/3609326>

- EDWARDS, Rochelle – KORTUM, Philip: He Says, She Says: Does Voice Affect Usability? 56(1) *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2012) 1486–1490.
<https://doi.org/10.1177/1071181312561295>
- FETTEROLF, Elizabeth – HERTOOG, Ekaterina: It's Not Her Fault: Trust Through Anthropomorphism among Young Adult Amazon Alexa Users. *Convergence* (2023).
<https://doi.org/10.1177/13548565231200337>
- GÁLIK Mihály – CSORDÁS Tamás (szerk.): *A média gazdaságtanának kézikönyve*. Budapest, Médiatudományi Intézet, 2020.
- GALLEGO, J. Ignacio: The Value of Sound: Datafication of the Sound Industries in the Age of Surveillance and Platform Capitalism. 26(7) *First Monday* (2021).
<https://doi.org/10.5210/fm.v26i7.10302>
- GATES, Bill: *The Road Ahead*. London, Viking, 1995.
- HUMPHRY, Justine – CHESHER, Chris: Preparing for Smart Voice Assistants: Cultural Histories and Media Innovations. 23(7) *New Media & Society* (2021) 1971–1988.
<https://doi.org/10.1177/1461444820923679>
- KÖVECSES, Zoltán: *Metaphor: A Practical Introduction*. Oxford, Oxford University Press, 2010.
- LEE, Chaiwoo et al.: User Acceptance of Voice Interfaces in the Automobile. 59(1) *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2015) 1641–1645.
<https://doi.org/10.1177/1541931215591355>
- LIND, Miriam – DICHEL, Sascha: Speaking, but Having No Voice: Negotiating Agency in Advertisements for Intelligent Personal Assistants. 30(3) *Convergence* (2023) 1008–1024.
<https://doi.org/10.1177/13548565231192100>
- MANOVICH, Lev: *The Language of New Media*. Cambridge, MIT, 2001.
<https://doi.org/10.22230/cjc.2002v27n1a1280>
- MARVIN, Carolyn: *When Old Technologies Were New Thinking About Electric Communication in the Late Nineteenth Century*. Oxford, Oxford University Press, 1990.
<https://doi.org/10.1093/oso/9780195063417.001.0001>
- MINDER, Bettina et al.: Voice Assistants in Private Households: A Conceptual Framework for Future Research in an Interdisciplinary Field. 10 *Humanities and Social Sciences Communications* 173 (2023).
<https://doi.org/10.1057/s41599-023-01615-z>
- MIN LEE, Kwan et al.: Voice Orientation of Conversational Interfaces in Vehicles. 43(3) *Behaviour & Information Technology* (2023) 433–444.
<https://doi.org/10.1080/0144929X.2023.2166870>
- NATALE, Simone: The ELIZA Effect: Joseph Weizenbaum and the Emergence of Chatbots. In Simone NATALE: *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford, Oxford University Press, 2021.
<https://doi.org/10.1093/oso/9780190080365.001.0001>
- NATALE, Simone – BORY, Paolo – BALBI, Gabriele: The Rise of Corporational Determinism: Digital Media Corporations and Narratives of Media Change. 36(4) *Critical Studies in Media Communication* (2019) 323–338.
<https://doi.org/10.1080/15295036.2019.1632469>

- NATALE, Simone – COOKE, Henry: Browsing with Alexa: Interrogating the Impact of Voice Assistants as Web Interfaces. 43(6) *Media, Culture & Society* (2021) 1000–1016.
<https://doi.org/10.1177/0163443720983295>
- NGUYEN, Quynh N. – TA, Anh – PRYBUTOK, Victor: An Integrated Model of Voice-User Interface Continuance Intention: The Gender Effect. 35(15) *International Journal of Human-Computer Interaction* (2019) 1362–1377.
<https://doi.org/10.1080/10447318.2018.1525023>
- PARISI, David: *Archaeologies of Touch: Interfacing with Haptics from Electricity to Computing*. Minneapolis, University of Minnesota Press, 2018.
<https://doi.org/10.5749/j.ctt20mvgvz>
- PHILIP, George – SMITH, F. J. – CROOKES, Danny: Voice Input/Output Interface for Online Searching: Some Design and Human Factor Considerations. 14(2) *Journal of Information Science* (1988) 93–98.
<https://doi.org/10.1177/016555158801400204>
- PINCH, Trevor – BIJKER, Wiebe E.: Tények és termékek társadalmi konstrukciója, avagy hogyan segítheti egymást a tudomány szociológiája és a technika szociológiája (ford. KIRÁLY Gábor). *Replika*, 51–52. (2005) 57–87.
- RAKOTOMALALA, Francis et al.: Voice User Interface: Literature Review, Challenges and Future Directions. 1(2) *System Theory, Control and Computer Journal* (2021) 65–89.
<https://doi.org/10.52846/stccj.2021.1.2.26>
- SEITER, Ellen et al. (szerk.): *Remote Control: Television, Audiences, and Cultural Power*. London, Routledge, 1989.
- SIMON, Felix M.: Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy. 10(10) *Digital Journalism* (2022) 1832–1854.
<https://doi.org/10.1080/21670811.2022.2063150>
- SUTTON, Selina J. et al.: Voice as a Design Material: Sociophonetic Inspired Design Strategies in Human-Computer Interaction. In *CHI Conference on Human Factors in Computing Systems Proceedings*, 2019, 1–14.
<https://doi.org/10.1145/3290605.3300833>
- TÓFALVY Tamás: Meta-metaverzum: Egy technológia diszkurzív megalkotása, a platformizáció logikája és a jövő kérdőjelei. *Szabad Piac*, 2023/1., 53–62.
- TSIMHONI, Omer – SMIT, Daniel – GREEN, Paul: Address Entry While Driving: Speech Recognition Versus a Touch-Screen Keyboard. 46(4) *Human Factors* (2004) 600–610.
<https://doi.org/10.1518/hfes.46.4.600.56813>
- VLAHOS, James: *Talk To Me: How Voice Computing Will Transform the Way We Live, Work, and Think*. New York, Harper, 2019.
- WYATT, Sally: Metaphors in Critical Internet and Digital Media Studies. 23(2) *New Media & Society* (2021) 406–416.
<https://doi.org/10.1177/1461444820929324>