Médiakutató, 2026. tél
XXVII. évf. 4. szám, 45–53. o.
10.55395/MK.2026.4.4

Úgy tanul, ha kérdez...

Interpretáció, hermeneutika és a nagy nyelvi modellek

Krasztev Péter
Budapesti Gazdaságtudományi Egyetem

E tanulmány a mesterséges intelligenciát az interpretáció filozófiai problémájának újragondolására használja fel. Amellett érvel, hogy nem az a döntő kérdés, hogy interpretál-e a mesterséges intelligencia, hanem az, hogy miért éljük meg interpretációként a vele folytatott párbeszédet akkor is, amikor a másik oldalon nincs emberi tudat. Max Weber, Marcel Mauss, Norbert Elias, Clifford Geertz, Hans-Georg Gadamer és a hermeneutikai történetírás néhány meghatározó felismerésére támaszkodva arra a következtetésre jut, hogy a mesterséges intelligencia nem helyettesíti az interpretációt, hanem olyan technológiai közeget hoz létre, amely új fényben mutatja meg a megértés emberi feltételeit.

Kulcsszavak: filozófiai hermeneutika, Gadamer, interpretáció, megértés, mesterséges intelligencia, nagy nyelvi modellek

Learning by Asking

Interpretation, Hermeneutics, and Large Language Models

This paper uses artificial intelligence as a point of departure for rethinking the philosophical problem of interpretation. It argues that the central question is not whether artificial intelligence is capable of interpretation, but why dialogue with large language models can be experienced as an interpretive event even in the absence of a human consciousness on the other side. Drawing on key insights from Max Weber, Marcel Mauss, Norbert Elias, Clifford Geertz, Hans-Georg Gadamer and the hermeneutic tradition in historiography, the article argues that artificial intelligence does not replace interpretation. Rather, it creates a new technological medium that casts fresh light on the human conditions of understanding.

Keywords: artificial intelligence, Gadamer, interpretation, large language models, philosophical hermeneutics, understanding

„A szellemtudományok logikája … a kérdés logikája."
(Gadamer 1984: 259)

Egy szerkesztői kérdés

Az itt következő gondolatmenetet egy szerkesztői üzenet indította el. Rövid, a tárgyra szorítkozó levél volt, amilyet naponta tucatjával kap és küld bárki, aki a szakmában dolgozik. Egyetlen kérdést tett fel: elfogadható-e tudományos folyóiratban olyan kézirat, amelynek elkészítésében a mesterséges intelligencia (AI) nemcsak közreműködött, hanem részben a cikk tárgya is?

Első ránézésre egyszerű technikai problémának tűnt, amelyre a szerkesztőség és a kutatóhelyek mind bevezettek már ilyen-olyan protokollt. A mesterséges intelligencia fejlődésének dinamikája azonban olyan, hogy – a klasszikus fordulattal élve – helyt- és időtálló választ találni körülbelül olyan vállalkozás, mint futás közben megborotválni egy nyulat. A vita jelen esetben is hamar a szokásos mederbe terelődött: mit szabad és mit nem, hol húzódik a határ ember és gép között, mi fér bele és mi nem egy ilyen együttműködésbe.

Pedig lehet, hogy a szerkesztő valójában nem is erre a kérdésre kereste a választ. Pontosabban: olyan kérdést tett fel, amely szükségszerűen újabb kérdéseket szült. Ez így rendjén is való a társadalomtudományban: a jó kérdések itt nem szokták lezárni a vitákat, ellenben gyakran újakat indítanak el.

Az első ilyen kérdés rögtön a szerzőségre vonatkozott: miért okoz ma ekkora bizonytalanságot a szerző személye, amikor Roland Barthes (1968/1996) közel fél évszázada „kivégezte" a szerzőt, és azóta az irodalom- és a művészetelmélet, vagy éppen a posztmodern próza (mint például Jorge Luis Borges, Danilo Kiš vagy Esterházy Péter munkássága) számtalan formában relativizálta az egyéni szerzői tudat elsődlegességét? A műalkotás már rég nem kizárólag egy autonóm alkotó „terméke" vagy tulajdona, hanem szövegek, idézetek, kulturális mintázatok és korábbi értelmezések kusza hálózata – a művészetben ez ma már többnyire közhelynek is számít.

A társadalomtudomány azonban más helyzetben van. A tudományos szövegek is korábbi ismeretek újrarendezéséből, fogalmi keretek átalakításából és új értelmezésekből születnek. Mégis sokkal nehezebben mondunk le a szerző kitüntetett szerepéről. Ennek részben intézményi okai vannak: a tudományos életben a szerző neve nem pusztán egy szöveghez kapcsolódik, hanem hivatkozásokhoz, pályázatokhoz, előmenetelhez és szakmai tekintélyhez. A szerző relativizálása ezért itt nem csupán filozófiai kérdés, hanem a tudomány működésének egyik alapfeltételét érinti.

Éppen ezért talán nem is az a legfontosabb kérdés, hogy ki tekinthető egy mesterséges intelligencia segítségével készült tanulmány szerzőjének. Mielőtt erre rátérnék, érdemesebb tisztázni, mi történik akkor, amikor valaki órákon, napokon vagy akár heteken át párbeszédet folytat egy nagy nyelvi modellel. Minek nevezzük ezt a folyamatot? Információkeresésnek? Konzultációnak? Szerkesztésnek? Közös szövegalkotásnak? Vagy valami olyannak, amire egyelőre nincs pontos szavunk?

Az elmúlt néhány év mesterséges intelligenciáról szóló vitái sok energiát fordítottak annak eldöntésére, hogy a nagy nyelvi modellek „értenek-e", „gondolkodnak-e", „kreatívak-e", vagy éppen „interpretálnak-e". Mintha a vita végső tétje az lett volna, sikerült-e az emberi megismerés valamely jól körülhatárolható képességét algoritmizálni (Bender et al. 2021, Floridi 2023, Mitchell & Krakauer 2023). Az utóbbi években egyre több szerző vetette fel, hogy a nagy nyelvi modellekről folytatott viták nem csupán a válaszok, hanem már a kérdésfeltevés szintjén is újragondolásra szorulnak (Coeckelbergh & Gunkel 2024). Ez az írás ehhez a felismeréshez kapcsolódik, de nem dekonstruktív, hanem hermeneutikai irányból közelít: nem azt kérdezi, interpretál-e a mesterséges intelligencia, hanem azt, hogy milyen értelemben válhat interpretatív eseménnyé az ember és a nagy nyelvi modell közötti párbeszéd.

Ez természetesen a nagy nyelvi modellek használatának csak bizonyos helyzeteire érvényes. A modell szolgálhat információkeresésre, fordításra, összefoglalásra vagy számos más gyakorlati célra is. Interpretatív eseménnyé azok a tartósabb, dialogikus helyzetek válhatnak, amelyek során a felhasználó saját kérdéseit, előfeltevéseit és értelmezési lehetőségeit is újra meg újra játékba hozza. A továbbiakban elsősorban az AI-használatnak ezek a formái érdekelnek.

Interpretáció vagy mintázatfelismerés

Az interpretáció fogalmát gyakran magától értetődőbbnek tekintjük, mint amilyen valójában. Beszélünk szövegek, cselekvések, történelmi események vagy kulturális jelenségek értelmezéséről anélkül, hogy minden alkalommal tisztáznánk, mi különbözteti meg az interpretációt az információk rendszerezésétől, az oksági magyarázattól vagy a szabályszerűségek leírásától. Bár az interpretáció fogalmának tisztázására a hermeneutikában, a társadalomtudományban és az irodalomelméletben számos jelentős kísérlet született (Taylor 1971, Ricoeur 1976), a fogalom használata ma is feltűnően sokértelmű. Ez a bizonytalanság mindaddig inkább elméleti jelentőségűnek tűnhetett, amíg az összefüggő és meggyőző interpretációk létrehozása kizárólag emberi tevékenységnek látszott. A nagy nyelvi modellek megjelenése azonban éppen ezt a hallgatólagos bizonyosságot tette kérdésessé. Ha egy statisztikai valószínűségek alapján működő rendszer képes olyan szöveget létrehozni, amelyet értelmezésként olvasunk, akkor vagy a gép kezdett interpretálni, vagy nekünk kell pontosabban meghatároznunk, mit értünk interpretáció alatt, vagy – ahogy David Bordwell (1989/1991: 1) a filmek értelmezéséről írt monográfiájában megfogalmazza –: „…legelső feladatunk tehát interpretálni az »interpretációt«".

A továbbiakban itt három, egymással összefüggő, de nem azonos értelemben merül fel az interpretáció fogalma. Interpretációnak nevezzük magát az értelmezés aktusát, az ennek eredményeként létrejövő szöveget, valamint azt az eseményt, amelyben egy kérdés vagy egy értelmezői horizont átalakul. Egy nagy nyelvi modell képes interpretációként olvasható szöveget létrehozni anélkül, hogy ebből önmagában következne, hogy interpretatív aktust végez. A jelen tanulmány szempontjából elsősorban a fogalom harmadik jelentése érdekes: miként válhat az ember és a modell közötti párbeszéd interpretatív eseménnyé.

A nagy nyelvi modell működése leegyszerűsítve korábban előfordult nyelvi elemek és összefüggések valószínűségi feldolgozásaként írható le. Nem kész mondatokat keres elő, hanem a kontextus alapján hoz létre új szöveget. A mintázatok új kombinációi korábban nem létező összefüggéseket is létrehozhatnak. Az újdonság azonban önmagában még nem interpretáció. E különbség klasszikus megfogalmazása John R. Searle (1980) kínaiszoba-érve: a formális jelek szabályszerű kezelése önmagában nem hoz létre szemantikai tartalmat, ezért a helyes válaszokból még nem következik megértés. Daniel C. Dennett (1987) érvei sokkal érzékenyebb regiszterben mozognak: szerinte az intencionalitás nem feltétlenül valamely belső tulajdonság közvetlen felismerése, hanem olyan értelmezési beállítódás is lehet, amelyet a rendszer viselkedésében feltárható, magyarázó erejű mintázatok igazolnak. A mai nagy nyelvi modellekről (*Large Language Model,* röviden: LLM) folyó vitákban ugyanez az alapvető nézetkülönbség tér vissza, immár a jelentés, a reprezentáció és a megértés kérdéseire is kiterjedve (Millière & Buckner 2026). A vita ezen a ponton nem oldja meg, inkább kijelöli ennek a gondolatmenetnek a kérdését: van-e különbség aközött, hogy egy rendszer valóban megért, és aközött, hogy a vele folytatott párbeszéd interpretatív eseménnyé válik?

Ennek a kérdésnek a vizsgálatához a jelen kontextusban nem szükséges eldöntenünk, rendelkezhet-e egy nagy nyelvi modell tudattal, intencionalitással vagy a szó emberi értelmében vett megértéssel. A kortárs szakirodalomban magának a szemantikai megértésnek a megítélése is több, egymással versengő álláspontra bomlik, és olyan érvek is megjelentek, amelyek az LLM-eknek valamilyen értelemben szemantikát, sőt nyelvi megértést tulajdonítanak (Havlík 2025, Søgaard 2025). Ezeket az ontológiai kérdéseket ezért a továbbiakban zárójelben hagyom. A vizsgálat arra irányul, hogy mi történik az emberi értelmezéssel a modellel folytatott párbeszéd során, ehhez pedig nem kell ismernünk a modell belső állapotát.

A mintázatfelismerés önmagában még nem interpretáció, de a határvonalat aligha húzhatjuk meg a mechanikus ismétlés és az eredeti gondolat között. Az emberi gondolkodás is nagyrészt analógiák, felismerések és korábban elsajátított sémák újrarendezésére épül. Az interpretáció során azonban a felismert mintázat jelentőssé válik. Az értelmező kapcsolatot teremt az elemek között, és egyúttal azt állítja, hogy ez a kapcsolat teszi érthetőbbé azt a világot, amelyben az adott cselekvés, szöveg vagy esemény létrejött.

A társadalomtudósok közül Max Weber (1922/1987) megértő szociológiájának kiindulópontjában körvonalazódik ez a megközelítés. A társadalmi cselekvés nem attól társadalmi, hogy szabályszerűséget mutat, hanem attól, hogy a cselekvő azt szubjektív értelemmel kapcsolja össze, és mások viselkedéséhez igazodik. A társadalomtudomány ezért nem elégedhet meg a viselkedések leírásával, hanem azt is meg kell értenie, mit jelentenek azok a cselekvő

számára. Ezzel együtt a megértés Webernél nem empatikus beleélés, hanem módszertani rekonstrukció: annak vizsgálata, hogy egy cselekvés milyen jelentést nyer a rendelkezésre álló társadalmi kategóriák és lehetőségek között.

Marcel Mauss-nál (1925/2004b) is jól kitapintható ugyanez a logika: a kulturális jelentések a társadalmi természeten keresztül mutatkoznak meg. Mauss a *Tanulmány az ajándékról* című írásában kimutatja, hogy az ajándékcsere – a *potlatch* – jelentése nem a tárgyban vagy az adományozó szándékában rejlik, hanem a kapcsolatok, a kötelezettségek és a kollektív elképzelések rendszerében. Ugyanez a szervezőelv sejlik fel *A test technikái* című tanulmányában is: a legkézenfekvőbb testi gesztusok is társadalmilag tanult és történetileg változó technikák. Egy mozdulat értelmezéséhez ezért nem elegendő a formájának leírása, hanem ismernünk kell azt a társadalmi világot is, amelyben az természetesnek, helyénvalónak vagy sértőnek számít (Mauss 1936/2004a). Mauss ezzel a durkheimi társadalmi tény fogalmát radikalizálja: a „teljes társadalmi tény" már nem egyetlen társadalmi dimenzióra utal, hanem olyan jelenségre, amelyben gazdasági, jogi, vallási, erkölcsi és rokonsági viszonyok egyszerre szerveződnek egységgé.

Norbert Elias (1939/1987a) ugyanezt a gondolatot történeti léptékben alkalmazza. A civilizáció folyamata nem tények egymás mellé rendezése, hanem olyan értelmezés, amely a viselkedésformák, az intézmények és az érzékenységek között történeti kapcsolatot teremt. Más kérdésből kiindulva ugyanezek az adatok más történetet mondanának el.

A fentiekből távolról sem az következik, hogy minden interpretáció önkényes volna. Éppen azért vitatható, mert meg kell mutatnia, milyen tapasztalati anyagra épül, és a tények bemutatott elrendezése és egymáshoz való viszonyítása miért teszi a vizsgált jelenséget érthetőbbé más lehetséges olvasatoknál. A tényekből kirajzolódó mintázat önmagában még nem hordozza saját értelmezését; az interpretáció attól válik meggyőzővé, hogy megmutatja, milyen összefüggésben nyer értelmet a vizsgált jelenség.

Meddig sűríthető egy leírás?

Clifford Geertz (1973/1994) interpretív antropológiája a weberi kiindulópontot önálló kultúraelméleti és etnográfiai programmá alakította. A kultúrát nem az emberi viselkedést meghatározó szabályok összességeként, és nem is az egyének fejében tárolt ismeretek halmazaként fogta fel, hanem az emberek által szőtt jelentéshálók nyilvános világaként. Geertz Weberre hivatkozva, de saját metaforájával fogalmazza meg, hogy az ember „a jelentések maga szőtte hálójában függő állat", vagyis a kultúra elemzése nem törvényeket kereső tudomány, hanem jelentéseket kutató, interpretív vállalkozás (Geertz 1973/1994: 194–195). A hangsúly itt nemcsak a jelentésen, hanem annak nyilvános természetén van. Egy kacsintás azért értelmezhető, mert társadalmilag felismerhető gesztus – mondja Geertz. Ugyanaz a szemhéjmozdulat lehet önkéntelen rándulás, titkos jelzés vagy ezek paródiája; a fizikai esemény azonos marad, kulturális jelentése azonban megváltozik. A sűrű leírás feladata ezért nem a részletek felhalmozása, hanem annak a kontextusnak a rekonstruálása, amelyben valamely értelmezés érvényessé válik.

Ezen a ponton talán már tisztábban látható a mintázatfelismerés és az interpretáció közötti különbség. Egy nagy nyelvi modell felismerheti a kacsintás lehetséges jelentéseit, és a kontextus alapján valószínűségi sorrendet állíthat fel közöttük. Ez azonban még nem azonos azzal az interpretív művelettel, amelyben a kutató eldönti, hogy egy konkrét társadalmi világban mi történt, majd ezt a helyi jelentések rekonstruálásával teszi megindokolhatóvá.

Ugyanakkor Geertz számára az interpretáció nem a „szubjektív" módszer alternatívája. Az etnográfiai leírás konstrukció, de nem önkényes fikció. A kutató mások értelmezéseit értelmezi: azt próbálja rekonstruálni, miként értelmezik saját cselekvéseiket azok, akiknek a világáról ír. Az antropológiai szöveg ezért szükségképpen másod- vagy harmadrendű interpretáció, amelynek hitelessége nem a nézőpont hiányából, hanem reflektált és a vizsgált társadalmi világhoz visszakötött alkalmazásából ered. A sűrű leírást ezért félrevezető pusztán részletgazdag leírásként érteni. Nem attól sűrű, hogy sok információt tartalmaz, hanem attól, hogy feltárja a jelentés különböző szintjeinek kapcsolatát. Egy algoritmus elvileg nagyságrendekkel több adatot rendelhet egy jelenséghez, mint amennyit bármely kutató át tud tekinteni. A leírás mégis „vékony" maradhat, ha nem képes megmutatni, hogy az adatok közül melyek alkotják a résztvevők számára értelemmel bíró világot. A sűrűség nem mennyiségi tulajdonság, hanem interpretív teljesítmény.

Az AI mint hermeneutikai közeg

Ha azonban az interpretáció valóban így működik, a nagy nyelvi modelleknek is lehet szerepük benne. Páratlan sebességgel képesek lehetséges összefüggéseket létrehozni, párhuzamokat keresni, fogalmi változatokat kipróbálni, ellenpéldákat kínálni és egy kérdést különböző tudományos hagyományok felől újrafogalmazni. Egy kutatási probléma köré így tények, fogalmak, analógiák és lehetséges magyarázatok széles köre rendezhető, majd a következő kérdéssel újra szűkíthető vagy más irányból közelíthető meg. A hagyományos könyvtári vagy az internetes kereséshez képest talán a „zoomolás" sebessége jelenti a legfontosabb változást: ugyanaz a probléma néhány kérdésváltás alatt újabb és újabb fogalmi léptékben és összefüggésben jelenhet meg. Az interpretáció lehetséges útjai ezzel gyorsabban elágaznak, és olyan kapcsolatok is láthatóvá válhatnak, amelyek az értelmező számára addig nem voltak hozzáférhetők.

A modell ezzel interpretációs lehetőségek sokaságát kínálja fel, nyersanyagot az értelmezéshez. Ezek közül valakinek ki kell választania azokat, amelyekért elméletileg és empirikusan, valamint – különösen más emberek világának leírásakor – empatikusan és erkölcsileg is felelősséget tud vállalni. Ez a felelősség nem külső etikai követelmény, hanem az interpretáció szerkezetéből fakad. Az értelmező a lehetséges kapcsolatok között súlyoz, és azt állítja, hogy bizonyos kapcsolatok jelentősebbek másoknál. Az interpretáció így túllép a mintázatok gyakoriságán, és arra kérdez rá, mi az, ami jelentőséggel bír – akkor és ott.

A lehetséges összefüggések megsokszorozódásával még fontosabbá válik az a feszültség, amelyet Elias (1983/1987b) a „bevonódás és távolságtartás" egyensúlyaként ír le. A társadalomtudós nem állhat teljesen kívül az általa vizsgált világon, mert fogalmai és kérdései maguk is társadalmi tapasztalatokból fakadnak. Ugyanakkor nem mondhat le a távolságtartásról sem, mert enélkül nem tudna kilépni a saját előfeltevéseiből. Az interpretáció tehát nem a bevonódás ellentéte, hanem a kettő folyamatos egyensúlya.

A nagy nyelvi modell helyzete ebből a szempontból más. Elias bevonódásfogalma olyan tapasztalati és társadalmi pozíciót feltételez, amelyből az értelmező reflexív módon távolságot teremthet. A modell működésének leírásakor ilyen pozíciót nem szükséges feltételeznünk. Az AI betanítása során felhasznált szöveganyag történeti és kulturális torzításai ettől még megjelenhetnek a válaszaiban, és az önreflexió nyelvi formáit is képes újraalkotni: figyelmeztethet saját korlátaira, vagy alternatív nézőpontokat kínálhat. Mindez azonban önmagában nem dönti el, hogy ezek mögött a nyelvi műveletek mögött áll-e olyan tapasztalati viszony önmaga és a világ között, amelyről Elias fogalmai szólnak.

Hasonló irányból fogalmazta meg kritikáját Hubert Dreyfus (1972/1992) is, aki szerint az emberi megértés nem formális szabályok alkalmazásából, hanem egy testileg és társadalmilag megélt „háttérvilágból" bontakozik ki. Nem elsősorban információkat dolgozunk fel, hanem eleve olyan világban tájékozódunk, amelyben bizonyos dolgok jelentőséggel bírnak, mások pedig háttérben maradnak. A nagy nyelvi modellek sikere ugyanakkor arra utal, hogy ennek a háttérvilágnak a nyelvi lenyomatai statisztikailag jóval gazdagabban rekonstruálhatók, mint amire a Dreyfus által bírált korabeli AI-rendszerek képesek voltak. Hasonló különbséget tesz Christoph Durt (2025) a jelentést hordozó tágabb kontextus és az LLM-ek által feldolgozott szöveges *co-text* között. A nyelvi lenyomatok egyre meggyőzőbb rekonstrukciója ugyanakkor nyitva hagyja, milyen viszonyban áll mindez magával a Dreyfus által leírt háttérmegértéssel.

Ettől persze még az AI és az ember viszonya nem írható le egyszerűen úgy, hogy az egyik mintázatokat dolgoz fel, a másik pedig interpretál. Az emberi interpretáció is nyelvi mintázatokra, tanult sémákra és korábbi értelmezésekre épül. Eredetisége gyakran nem új elemek létrehozásában, hanem a meglévők új kérdés szerinti újrarendezésében áll. A nyelvi modell pedig részt vehet ebben a folyamatban. És nem feltétlenül születik belőle kétféle szöveg. Egy AI által létrehozott bekezdés nyelvi megformáltsága önmagában nem árulja el, milyen interpretatív folyamatban jött létre, hiszen ember is írhat pusztán bevett mintázatokat ismétlő szöveget. Az interpretatív eseményt ezért nem a szöveg eredete határozza meg, hanem az a folyamat, amelyben valaki egy lehetséges összefüggést saját kérdéséhez, tapasztalati anyagához és a megérteni kívánt világhoz mér.

Innen nézve a mesterséges intelligenciával folytatott párbeszéd sajátos interpretív helyzetet hoz létre. A modell egymás mellé rendez, variál és újrafogalmaz; a felhasználó felismer, elvet, korrigál és hangsúlyokat helyez át. A párbeszéd nemcsak a válaszokat, hanem magát a kérdést is megváltoztathatja, és ezzel azt is, ami a vizsgált anyagból láthatóvá válik. Az AI tehát nem azért teszi újra aktuálissá Geertzet, mert maga is interpretál. Éppen

ellenkezőleg: arra figyelmeztet, hogy az interpretáció mindig kérdések, válaszok, szövegek és korábbi értelmezések közötti mozgásban jött létre. A nagy nyelvi modell ebbe a térbe lép be: nem jelentésteremtő szubjektumként, hanem dialogikus közegként, amely rendkívüli sebességgel rendezi át a lehetséges kapcsolatok mezőjét. Az AI tehát nem Geertz antropológusa, ugyanakkor a vele dolgozó kutató könnyen geertzi helyzetben találja magát. Már eleve megformált, kulturálisan rétegzett jelentések között mozog, és ezek közül kell kiválasztania azokat a kapcsolatokat, amelyek saját kérdésének perspektívájából valóban mondanak valamit a vizsgált világról. A technológia újdonsága ezért nemcsak, hogy nem teszi feleslegessé az interpretív hagyományt, hanem éppenséggel láthatóvá teszi egyik alapvető felismerését. Az adat, a szöveg és a mintázat önmagában nem beszél. Csak valamely kérdésre válaszolva kezd el jelentéssel bírni.

Horizontok találkozása egy algoritmussal

Első pillantásra triviálisnak tűnhet, hogy minden válasz valamely kérdésre adott reakció. Az interpretáció szempontjából azonban a kérdés nem egyszerűen megelőzi a választ, hanem megteremti annak feltételét. Nemcsak azt dönti el, mire keressük a választ, hanem azt is, mi válhat egyáltalán releváns ténnyé vagy értelmezhető összefüggéssé. Ugyanaz a szöveg, történeti forrás vagy társadalmi jelenség más kérdéssel megközelítve eltérő jelentést nyerhet, miközben maga az empirikus anyag változatlan marad. Az interpretáció újdonsága ezért gyakran nem új adatok felfedezésében, hanem a kérdés átalakulásában rejlik.

A filozófiai hermeneutika nem új módszert kínál fel, hanem abból a felismerésből indul ki, hogy a megértés maga sem írható le puszta módszerként. Hans-Georg Gadamer (1960/1984) szerint a megértés esemény, amelyben az értelmező maga is részt vesz. Nem semleges megfigyelőként közelít a világhoz, hanem történeti és nyelvi horizontjából. Fogalmai, kérdései és érzékenysége ebből a világból származnak; a megértés mindig már meglévő jelentések között bontakozik ki. A kérdés ezért nem egyszerű információigény, hanem kijelöli azt a horizontot, amelyben valami egyáltalán jelentést nyerhet. A valódi kérdés nem lezárja, hanem megnyitja az értelmezést.

Ez a felismerés ma már rég kilépett a filozófiai hermeneutika kizárólagos illetékességéből: számos más tudományterületen találunk vele rokon felismeréseket. A kognitív pszichológia mentális sémákról (Kahneman 2011/2013), a kognitív nyelvészet *frame*-ekről és konceptuális metaforákról (Lakoff 1987, Lakoff & Johnson 2003), Thomas Kuhn (1962/2000) paradigmákról, Polányi Mihály (1966/1997) hallgatólagos tudásról beszél. Ezek a sok szempontból különböző elméleti hagyományok más-más összefüggésben és fogalmi apparátussal ugyan, de egyaránt számolnak azzal, hogy a megértés előzetes orientációkból indul ki. Gadamer eredetileg 1960-ban megjelent saját hozzájárulása nem ennek felismerése, hanem a filozófiai következményeinek levonása. Amit a felvilágosodás leküzdendő előítéletnek tekintett, azt ő a megértés feltételének nevezi. Az előítélet *(Vorurteil)* nem irracionális torzítás, hanem történetileg kialakult előzetes ítélet, amely nélkül nem tudnánk kérdezni, különbséget tenni fontos és lényegtelen között, vagy értelmes egészként olvasni a világot. A megértés feladata ezért nem az előítéletek felszámolása, hanem azoknak folyamatos próbára tétele (Gadamer 1960/1984).

A megértés így nemcsak a vizsgált jelenséget, hanem magát az értelmezőt is alakítja. Gadamer (1960/1984) ezt nevezi a horizontok összeolvadásának *(Horizontverschmelzung)*. A horizont nem egyszerű nézőpont, hanem annak történetileg kialakult lehetőségtere, amely meghatározza, mit tartunk kérdezhetőnek vagy jelentősnek. A megértés során ezért nemcsak a vizsgált jelenség értelmezése, hanem a saját horizontunk is módosul. Ez a változás azonban nem az értelmező szubjektivitásának diadala. Gadamer éppen azt igyekszik meghaladni, hogy a megértést a szubjektum belső aktusaként fogjuk fel. Az interpretáció nem azért személyes, mert kizárólag az értelmező tudatában zajlik, hanem mert történetileg mindig valaki ért meg valamit valahonnan. A szubjektivitás ezért nem az interpretáció forrása, hanem egyik feltétele: nem önkényes nézőpontot, hanem történeti elhelyezkedést jelent. A megértés hitelessége nem azon múlik, hogy sikerül-e ettől megszabadulni, hanem azon, hogy az értelmező képes-e saját horizontját is a megértés tárgyává tenni. A hermeneutikai szubjektum fogalma ezzel maga is átalakul. Az értelmező nem a jelentés kizárólagos létrehozója, de nem is passzív befogadója. A jelentés a megértés eseményében jön létre, amely egyszerre alakítja a vizsgált jelenség értelmezését és az értelmező önértelmezését. A másik megértése ezért mindig saját előfeltevéseink átalakulása is.

Ez a gondolat különösen termékennyé válik a nagy nyelvi modellek értelmezésekor. Az AI-val folytatott párbeszéd során ugyanis könnyen úgy tűnhet, mintha a modell maga is részt venne ebben a hermeneutikai eseményben. Válaszai új összefüggéseket tárnak fel, váratlan fogalmi kapcsolatokat javasolnak, korábbi előfeltevéseinket megkérdőjelezik, és nem ritkán magát a kérdést is átalakítják. A felhasználó gyakran valóban úgy éli meg, hogy a párbeszéd során saját horizontja módosul. A horizontok összeolvadását azonban szigorú értelemben nem szükséges a felhasználó és a modell közötti szimmetrikus viszonyként elképzelnünk. A gadameri horizont történeti tapasztalatot és önértelmezést feltételez; a modell működésének leírásához itt ilyen saját horizontot nem kell tételeznünk. A válaszaiban ugyanakkor az AI betanítása során felhasznált szöveganyag emberi értelmezéseinek statisztikai mintázatai rendeződnek újra. Ez nem zárja ki a horizontok összeolvadását: a modell olyan hagyományokat és értelmezési lehetőségeket közvetíthet, amelyek átalakítják az ember horizontját.

Ebben az értelemben a nagy nyelvi modell sajátos hermeneutikai közeggé válik. Óriási sebességgel teszi hozzáférhetővé és rendezi újra azt az interpretációs hagyományt, amelyen minden emberi megértés eleve nyugszik. Nem helyettesíti az interpretációt, hanem láthatóvá teszi annak egyik alapvető sajátosságát: hogy minden értelmezés korábbi értelmezésekkel folytatott párbeszédből születik.

A mesterséges intelligencia filozófiai jelentősége ezért túlmutat azon a kérdésen, hogy rendelkezik-e szubjektivitással vagy képes-e interpretálni: megjelenése új helyzetbe állítja a hermeneutikát. Gadamer felismerései ebben a helyzetben új próba elé kerülnek: pontosabban kell megfogalmaznunk, mely mozzanatok nélkülözhetetlenek a megértés eseményéhez, és melyeket kötöttünk eddig kizárólag az emberi szubjektumhoz.

Az objektivitás új jelentése

Ha a történetírás valóban a kérdések története, akkor óhatatlanul felmerül a klasszikus ellenvetés: mi marad így az objektivitásból? Nem vezet-e a hermeneutikai megközelítés relativizmushoz, ha minden korszak más kérdéseket tesz fel ugyanannak a múltnak?

Ez az aggály abból a pozitivista előfeltevésből indul ki, hogy az objektivitás az értelmező eltűnését jelenti. Mintha a történész feladata az volna, hogy megszabaduljon saját korától és előfeltevéseitől, s valamiféle történelmen kívüli nézőpontból szemlélje a múltat. A huszadik század tudományelmélete azonban éppen ezt a feltevést kérdőjelezte meg. Kuhn óta aligha tartható fenn az a meggyőződés, hogy a tudomány előfeltevések nélküli megfigyelések felhalmozásából épül fel. A tudományos közösségek mindig meghatározott fogalmi keretek között dolgoznak; ha ez a természettudományokra is igaz – mint azt Kuhn állítja –, még kevésbé várható el a történésztől, hogy függetlenítse magát saját történeti helyzetétől.

A hermeneutikai hagyomány azonban nem az objektivitás feladását, hanem annak újrafogalmazását jelenti. A történész nem azért közelíthet a múlthoz, mert képes felfüggeszteni saját előfeltevéseit, hanem mert képes reflektálni rájuk. Az önreflexió nem az objektivitás alternatívája, hanem egyik feltétele. A történeti objektivitás ezért nem nézőpont nélküli megismerést jelent, hanem azt a képességet, hogy az értelmező saját kérdéseit is vizsgálat tárgyává teszi. Nemcsak azt kérdezi, mit mondanak a források, hanem azt is, miért éppen ezt kérdezi tőlük. Saját horizontja így nem akadálya, hanem feltétele a megértésnek.

Talán ezért maradnak velünk a nagy történeti munkák akkor is, amikor számos konkrét megállapításuk korrekcióra szorul. Az Annales-iskola klasszikusainak (például Marc Bloch, Georges Duby és François Furet), illetve a hozzájuk kapcsolódó későbbi történetírás olyan szerzőinek, mint Robert Darnton a műveit főleg azért olvassuk máig, mert koruk alapvető kérdéseit ismerték fel, és ezek segítségével tették újra kérdezhetővé a múltat, még ha egyes következtetéseik azóta korrekcióra is szorultak. Munkáik egyszerre dokumentumai annak a kornak, amelyről írnak, és annak is, amelyben megszülettek. A történeti objektivitás ezért nem a szerző eltűnésében rejlik, hanem abban, hogy az értelmező saját történeti helyzetét is bevonja az interpretációba. Amíg saját horizontunk láthatatlan marad számunkra, addig nem megszűnik hatni, hanem észrevétlenül irányítja kérdéseinket és válaszainkat.

A mesterséges intelligencia korában ez a felismerés új dimenziót kap. Az önreflexiónak ma már nemcsak az értelmező előfeltevéseire kell kiterjednie, hanem arra a technológiai közegre is, amely egyre inkább részt vesz

az interpretációban. A kérdés ezért már nem csupán az, milyen előfeltevései vannak a történésznek, hanem az is, milyen szelekciós logikák és valószínűségi struktúrák alakítják annak a technológiai közegnek a működését, amelyben az interpretáció egyre gyakrabban végbemegy.

Zárszó helyett

A tanulmány elején abból a közkeletű kérdésből indultunk ki, képes-e a mesterséges intelligencia interpretálni. Mire azonban idáig eljutottunk, talán már világosabbá vált, miért érdemes másképp feltenni a kérdést.

Gadamer az *Igazság és módszer* második kiadásának előszavában úgy fogalmaz, hogy filozófiai kérdése nem az, mit teszünk és mit kellene tennünk, hanem az, hogy „akarásunkon és tevékenységünkön túl mi történik velünk" (Gadamer 1960/1984: 12). Talán a mesterséges intelligencia filozófiai jelentősége is itt ragadható meg. A valódi kérdés nem az, hogy a modell rendelkezik-e szubjektivitással, vagy milyen feladatokat képes elvégezni, hanem az, mi történik a megértéssel, amikor az interpretáció új technológiai közegben megy végbe.

Jorge Luis Borges (1945/2015) *Az Alef* című novellájának narrátora egyetlen pontban pillantja meg a világ teljességét. A látomás azonban nem zárja le a keresést. A novella végén már azon töpreng, vajon valóban az Alefet látta-e, vagy csupán egy másik, tökéletlen mását. A tapasztalat nem bizonyosságot hagy maga után, hanem megváltozott kérdéseket. Már nem ugyanazt akarja tudni, mint mielőtt lenézett a pince mélyére.

Talán a mesterséges intelligenciával folytatott párbeszéd sem azért jelent fordulópontot az interpretáció történetében, mert választ ad egy régi kérdésre. Inkább azért, mert újra láthatóvá teszi, hogy a megértés soha nem pusztán válaszok keresése volt. A valódi interpretációban nemcsak az változik meg, amit megértünk, hanem az is, ahonnan kérdezünk. A nagy nyelvi modellek nem helyettesítik ezt a folyamatot. A velük folytatott párbeszédben azonban saját előfeltevéseink, fogalmaink és kérdéseink is újra a megértés tárgyává válhatnak.

És éppen ez teszi újra időszerűvé a hermeneutikát. Nem azért, mert a mesterséges intelligencia megválaszolja régi kérdéseinket, hanem mert – Borges Alefjéhez hasonlóan – mire elfordítjuk róla a tekintetünket, már nem ugyanazok vagyunk, mint akik először rápillantottunk.

Felhasznált irodalom

Barthes, Roland (1968/1996): A szerző halála. In: Roland Barthes (szerk.): *A szöveg öröme. Válogatott írások*, 50–55. o. Osiris Kiadó.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell (2021): On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FAccT'21 Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623. Association for Computing Machinery, https://doi.org/10.1145/3442188.3445922

Bordwell, David (1989/1991): *Making Meaning. Inference and Rhetoric in the Interpretation of Cinema*. Harvard University Press.

Borges, Jorge Luis (1945/2015): Az Alef. In: M. Nagy Miklós (szerk.): *Az első magyar költőhöz. Válogatott művek*, 224–241. o. Európa Könyvkiadó.

Coeckelbergh, Mark & David J. Gunkel (2024): ChatGPT: Deconstructing the Debate and Moving It Forward. *AI & Society*, vol. 39, pp. 2221–2231, https://doi.org/10.1007/s00146-023-01710-4

Dennett, Daniel C. (1987): *The Intentional Stance*. MIT Press.

Dreyfus, Hubert L. (1972/1992): *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Press.

Durt, Christoph (2025): The Meaning of Context and Co-Text for Human Understanding and Large Language Models. *Philosophy & Digitality*, vol. 2, no. 1, https://doi.org/10.18716/pd.v2i1.11666

Elias, Norbert (1939/1987a): *A civilizáció folyamata*. Gondolat Kiadó.

Elias, Norbert (1983/1987b): *Involvement and Detachment*. Basil Blackwell.

Floridi, Luciano (2023): AI as Agency without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models. *Philosophy & Technology*, vol. 36, no. 1, art. 15, https://doi.org/10.1007/s13347-023-00621-y

Gadamer, Hans-Georg (1960/1984): *Igazság és módszer*. Gondolat.

Geertz, Clifford (1973/1994): Sűrű leírás. Út a kultúra értelmező elméletéhez. In: *Az értelmezés hatalma. Antropológiai írások*, 194–226. o. Századvég Kiadó.

Havlík, Vladimír (2025): Meaning and Understanding in Large Language Models. *Synthese*, vol. 205, no. 1, art. 9, https://doi.org/10.1007/s11229-024-04878-4

Kahneman, Daniel (2011/2013): *Gyors és lassú gondolkodás*. HVG Könyvek.

Kuhn, Thomas S. (1962/2000): *A tudományos forradalmak szerkezete*. Osiris Kiadó.

Lakoff, George (1987): *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press, https://doi.org/10.7208/chicago/9780226471013.001.0001

Lakoff, George & Mark Johnson (1980/2003): *Metaforák a mindennapi életben*. Typotext Kiadó.

Mauss, Marcel (1936/2004a): A test technikái. In: Marcel Mauss (szerk.): *Szociológia és antropológia*, 425–459. o. Osiris.

Mauss, Marcel (1925/2004b): Tanulmány az ajándékról. Az ajándékcsere formája és értelme az archaikus társadalmakban. In: Marcel Mauss (szerk.): *Szociológia és antropológia*, 195–342. o. Osiris Kiadó.

Millière, Raphaël & Cameron Buckner (2026): The Philosophy of Language Models. *Philosophy Compass*, vol. 21, no. 3, e70095, https://doi.org/10.1111/phc3.70095

Mitchell, Melanie & David C. Krakauer (2023): The Debate over Understanding in AI's Large Language Models. *Proceedings of the National Academy of Sciences*, vol. 120, no. 13, e2215907120, https://doi.org/10.1073/pnas.2215907120

Polányi Mihály (1966/1997): A hallgatólagos dimenzió. In: Polányi Mihány (szerk.): *Tudomány és ember*, 163–241. o. Argumentum Kiadó & Polányi Mihály Szabadelvű Filozófiai Társaság.

Ricoeur, Paul (1976): *Interpretation Theory: Discourse and the Surplus of Meaning*. Texas Christian University Press.

Searle, John R. (1980): Minds, Brains, and Programs. *Behavioral and Brain Sciences*, vol. 3, no. 3, pp. 417–457, https://doi.org/10.1017/S0140525X00005756

Søgaard, Anders (2025): Do Language Models Have Semantics? On the Five Standard Positions. In: Wanxiang Che, Joyce Nabende, Ekaterina Shutova & Mohammad Taher Pilehvar (eds.): *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 25910–25922. Association for Computational Linguistics, https://doi.org/10.18653/v1/2025.acl-long.1258

Taylor, Charles (1971): Interpretation and the Sciences of Man. *The Review of Metaphysics*, vol. 25, no. 1, pp. 3–51.

Weber, Max (1922/1987): *Gazdaság és társadalom I. A megértő szociológia alapvonalai*. Közgazdasági és Jogi Könyvkiadó.