

„AI A PIACON”: MESTERSÉGES INTELLIGENCIA RENDSZEREK ADOPTÁCIÓJA MAGYAR ÁLLÁSHIRDETÉSEK TÖMEGES FELDOLGOZÁSA TÜKRÉBEN¹

SZABADOS LEVENTE

Frankfurt School of Finance and Management; Soproni Egyetem

A mesterséges intelligencia (AI) rendszerek gazdasági térnyerése átalakítja a munkaerőpiacot, de Magyarországon az adaptáció mértéke kevésbé ismert. Kutatásunkban több mint 1.3 millió, 2019–2024 között megjelent online állás-hirdetést elemeztünk természetesnyelv-feldolgozási (NLP) módszerekkel. Kétféle osztályozást alkalmaztunk: kulcsszavas (RegEx) keresést és egy nagy nyelvi modellre (LLM) épülő, DSPy-optimalizált szemantikus eljárást.

Eredményeink szerint az AI-val érdemben összefüggő állás-hirdetések aránya mindössze 1.34%, túlnyomórészt Budapesthez kötődnek, és elsősorban az IT, tanácsadás, pénzügyi szolgáltatások és autóipar területein jelennek meg. Azonosítottunk három jellegzetes vállalati stratégiát: (1) egyszeri, rövid időszakra koncentrált AI-felvételi hullámot, (2) folyamatos, stabil AI-munkaerő-építést, és (3) olyan „branding” stratégiát, amely minden hirdetésben említi az AI-t, függetlenül annak tényleges relevanciájától. A generatív AI technológiák munkaerőpiaci jelenléte jelenleg marginális.

1 Kontextus

A Mesterséges Intelligencia (Artificial Intelligence, továbbiakban AI) rendszerek fejlődése az utóbbi másfél évtizedben jelentős lendületet kapott, különösen a mély neurális hálós technológia, azaz *Deep Learning* (fogalom eredetéhez lásd még: Dechter, 1986, Fradkov, 2020) legújabb hullámának tekinthető úgynevezett *megalapozó modellek* (*foundational models*; Bommasani és tsai., 2022), ezen belül is a *nagy nyelvi modellek* (*Large Language Models*, LLM-ek; Brown és tsai., 2020) széles körű és alacsony költségű elérhetőségével (LLM-használat fogyasztói árának csökkenése tekintetében lásd még: Appenzeller, 2024).

A Mesterséges Intelligencia koncepciója azonban egyáltalán nem új keletű: az AI fejlődését több szakaszra oszthatjuk, melyek során a technológia eltérő formákban és megközelítésekkel kísérte meg a humán gondolkodás egyes aspektusainak szimulálását. Az 1950-es években megjelent AI rendszerek elsősorban szimbolikus logikán alapultak, melyek az *expert system* („szakértői rendszer”) paradigmába illeszkedtek: szigorú szabályrendszerekkel dolgoztak,

¹Beérkezett 2025. augusztus 16. DOI: <https://doi.org/10.15170/SZIGMA.56.1299>. E-mail: l.szabados@int.fs.de.

kézi munkával kodifikált tudásbázisokra építve. Ezek a rendszerek bizonyos környezetekben, mint orvosi diagnosztika vagy pénzügyi elemzések, működőképesek voltak, de nem tudtak dinamikusan alkalmazkodni új helyzetekhez.

A 2000-es évek elején a *Machine Learning* (ML) paradigmája vette át a vezető szerepet, amely lehetővé tette az adatokból történő tanulást explicit szabályok kódolása nélkül. Az ML módszerek, különösen a felügyelt tanulás, sikeresen alkalmazásra kerültek például képfelismerésben és piaci predikciókban. A technológia továbbfejlődése 2012-ben ért el egy áttörési pontot, amikor a mély neurális hálók (*Deep Learning*) igazolták dominanciájukat a számítógépes látás és egyéb AI alkalmazások terén. Ennek ellenére az AI ekkor még mindig szigorúan *narrow AI* maradt, azaz egyedi, célspecifikus modellekkel operált, melyek egy-egy meghatározott feladatra voltak optimalizálva.

A 2020 utáni időszakban az AI kutatás egy újabb paradigmaváltáson ment keresztül, amelyben elsősorban a *nagy nyelvi modellek* (LLM-ek) megjelenése jelentette az áttörést. Az LLM-ek nem csupán egy-egy konkrét feladatra lettek tréningezve, hanem általános problémamegoldási képességet mutattak, képesek voltak az emberi nyelv szintetizálására, és ezen keresztül komplex feladatok megoldására. Az OpenAI *ChatGPT* modelljének 2022. novemberi megjelenése széleskörűen demonstrálta e modellek képességét, ami globális médiavisszhangot és példa nélküli ipari és közgazdasági érdeklődést eredményezett.

Ennek fényében nem meglepő, hogy az AI munka világában történő alkalmazása is ugrásszerű fejlődésen ment keresztül (Anthropic, 2025), és a magyar üzleti gyakorlatban is megjelent (lásd még Szabados, 2025).

Amint azt korábbi cikkünkben (Szabados, 2024) részletesen elemeztük, a közgazdasági irodalomban is komoly témává vált az AI gazdasági, különösen munkaerőpiaci hatásának vizsgálata. Ebbe a kutatási trendbe kívánunk csatlakozni azáltal, hogy kvantitatív adatokon nyugvó elemzést adunk az AI magyarországi munkaerőpiaci helyzetéről.

1.1 Szakirodalmi kontextus

Anélkül, hogy a szakirodalom korábbi releváns elemzését (Szabados, 2024) megismételnénk, érdemes jelen vizsgálatunkat röviden elhelyezni a szakirodalom kontextusában.

A nemzetközi irodalom 2025-re egyre gazdagabb képet ad az AI munkaerőpiaci hatásairól: a munkakörök feladatszerkezete gyorsan változik, miközben a vállalatok többsége kifejezetten „augmentációs” (és nem tisztán létszámcsökkentő) felhasználást céloz. A World Economic Forum *Future of Jobs 2025* felmérése szerint a következő években a munkáltatók jelentős átalakulásra számítanak a tudásintenzív foglalkozásokban, különösen a generatív AI szélesebb integrációja miatt. (World Economic Forum, 2025) Ezzel párhuzamosan az európai szakpolitikai és mérési keretek is szilárdulnak: a JRC új, foglalkozási kitettséget mérő módszertana az AI-hoz való viszonyt feladatszintű térképezéssel ragadja meg, ami a gazdasági elemzések összeha-

sonlíthatóságát javítja (European Commission Joint Research Centre, 2025).

Empirikus szinten a friss, álláshirdetés-alapú vizsgálatok különösen relevánsak. A PwC 2025-ös Global AI Jobs Barometer közel egymilliárd hirdetés feldolgozásával kimutatja, hogy az AI-nak kitett foglalkozásokban nő a kereslet, és az AI-készségek a hirdetésekben szisztematikus bérprémiumhoz kapcsolódnak; a Lightcast 2025-ös szintézise pedig azt is dokumentálja, hogy az AI-készségek iránti igény már széles körben, a nem-tech ágazatokban is elterjedt (Lightcast, 2025; PwC, 2025). Közben az álláshirdetések tartalma is finomodik: az Indeed szerint az explicit „Responsible AI” kompetenciák gyorsan emelkedő, bár még kismértékű részarányt képviselnek az AI-hirdetéseken belül, ami az intézményi-etikai beágyazódás erősödésére utal (Gallacher & Aoki, 2025). Módszertanilag mindez alátámasztja az álláshirdetések szöveg-alapú, tartalmi (nem pusztán kulcsszavas) vizsgálatának szükségességét.

Az európai és közép-európai kontextusban kifejezetten fontos a regionális összevethetőség: a CEPS 2025-ös, 27 EU-tagállamra (és a tágabb európai térségre) kiterjedő hirdetés-adatbázison (Lightcast) nyugvó elemzése részletes képet ad az AI-készségek keresletének eloszlásáról ágazat és régió szerint, ami jó viszonyítási pont a magyar piac értelmezéséhez. (CEPS, 2025) Jelen munka – magyarországi online álláshirdetések 2019–2024 közti nagymintás, LLM-alapú szemantikus osztályozására építve – ehhez a nemzetközi kutatási irányhoz kapcsolódik, és a kulcsszavas említések zaját csökkentve a tartalmilag AI-releváns pozíciók feltárására törekszik.

1.2 Kutatásunk közvetlen kontextusa

(Ferenczy, 2025) megállapításai szerint a magyar AI nemzeti stratégia végrehajtásában jelentős megvalósítási rés (implementation gap) van: bár a stratégia kimondott céljai ambíciózusak, a hazai AI-fejlesztés és gazdasági hasznosulás szembetűnően alulmúlja a stratégiai lehetőségeket.

Az OECD jelentése (OECD, 2025) szintén erősen rámutat, 2023–24-ben felgyorsult a generatív AI üzleti bevezetése, de az úgynevezett „helyek, szektorok és vállalatok” közötti különbségek – térségi szinten is – jelentősek: a közép-európai régiók, köztük Magyarország, elmaradást mutatnak a fejlettebb nyugat-európai és amerikai területekhez képest, és noha nem kizárólag Magyarországra fókuszál, az OECD megállapításaiból levezethető, hogy Magyarország egyike azoknak a régióknak, amelyek lassúbb AI-adoptációval rendelkeznek.

2 Hipotézisek

- Magyarországon az AI adoptáció késik.
- Szektorálisan nincs egyensúlyban, néhány szektor kiemelkedik, többi elmarad.
- Területi egyenlőtlenségek állnak fenn.

További kutatási kérdés:

- Vajon a vállalatok AI-t adoptálnak, milyen stratégiákat követnek?

3 Módszertani megfontolások

Az AI megoldások munka világában történő alkalmazása egy piaci folyamat eredménye, amelynek egyik oldalán a technológia elérhetősége (ára, használatához szükséges tőkebefektetés mértéke), másik oldalán pedig annak hasznossága (például az AI megoldások által elérhető produktivitásnövekedés) áll. Emellett azonban bármiféle széles körű használat előfeltétele azon munkavállalók jelenléte egy szervezetben, akik érdemben képesek alkalmazni a technológiát, melyre azonban erős hatással lehet az AI-alapú munkalehetőségekről szóló tudás jelenléte, azaz egyáltalában az a tény, hogy egy munkavállaló találkozik olyan álláslehetőségekkel, melyekben lényeges elemet képez az AI képességek jelenléte. Ennek vizsgálatára – és egyben talán az AI alkalmazás egyfajta „proxyjaként” – véleményünk szerint igen alkalmas a nyilvános álláshirdetések szövegének elemzése. Munkánk során több mint 1.3 millió, Magyarország szempontjából releváns (munkavégzés helyét vagy hirdetés célzását tekintve Magyarországra célzott) online álláshirdetést elemeztünk 2019 és 2024 között számítógépes szövegelemzési módszerekkel.

3.1 Mit értünk „AI-alapú munkalehetőség” alatt?

Mivel az AI megoldások esetében, mint már említettük, általános célú technológiáról van szó, ami az üzleti folyamatok széles körében alkalmazható, pontos meghatározást igényel, mit értünk „AI-alapú munkalehetőség” alatt.

Jelen adatsor – sem időbeli fókusz (mivel később látni fogjuk, hogy a 2023-tól kezdődően elérhető még általánosabb, „generatív AI”-alapú megoldások elterjedése még várat magára), sem pedig részletessége (álláshirdetések, nem pedig részletes feladat- vagy tevékenységleírások) miatt – nem alkalmas arra, hogy olyan helyzetekről is számot adjon, melyben egy munkavállaló csupán érintőlegesen, felhasználói szinten, munkája során tulajdonképpen csak felszínesen használ valamely AI-alapú képességekkel is rendelkező informatikai megoldást (nevezzük egyszerűen „végfelhasználónak”), ezért vizsgálatunk során nem erre fókuszáltunk. Az alábbi elemzésben olyan munkalehetőségeket próbálunk azonosítani, melyeknek lényegi eleme az AI rendszerekkel való interakció, vagy azok fejlesztése, vagy célszerű, „fő fókusz” jellegű alkalmazása tekintetében. Ezt a függelékben található, osztályozáshoz használt LLM-prompt is jól illusztrálja, *melyben arra utasítjuk a modellt, hogy minden olyan esetet, ahol explicit módon említett bármilyen nyelven az alábbi technológiák legalább egyike, vagy azok jelenléte a szöveg kontextusából mint a munkavégzés explicit környezete/eszköze kikövetkeztethető, AI-alapú munkalehetőségnek tekintünk.* Technológiák: „Artificial Intelligence, Machine Learning, Data Science, Deep Learning, or related concepts such as Neural Networks, Natural Language Processing.”

3.2 Adatforrás

Az internetes tartalmak professzionális, nagy tömegű kinyerése jelentős üzleti területté vált, így például szűrt, és relevancia szerint előfeldolgozott álláshirdetések is „kulcsrakész” adatbázis formájában elérhetőek. Ezek közül mi a *JobsPikr* (Jobspikr, 2024) cég szolgáltatásait vettük igénybe, mely a fent említett időszámban és mennyiségben bocsátott áruba magyar, angol, és német nyelven összegyűjtött álláshirdetéseket, melyekre a gépi nyelvfeldolgozás eszköztárát alkalmazni tudtuk. Az adatbázis méretének és relevanciájának jó kontextust ad, ha mélyebben betekintünk a JobsPikr által publikált gyűjtési módszertanba. A szolgáltató célja, hogy nemzetközi összehasonlításban is stabil, kvázi „sztdardizált” adatsorokat állítson össze ügyfelei számára, melyek az eltérő országok és régiók trendjeit összevethetővé teszik – mely cél alapvetően pozitívnak tekinthető jelen elemzésünk szempontjából. A szolgáltató épp ezért 75000 fix, nagy lefedést biztosító internetes állásportált követ és dolgoz fel napi frissítéssel (és erős géptanulás-alapú szűréssel, utólagos osztályozással) mely lehetővé teszi a folyamatos változáskövetést. (Ez adatsorunkon is látható, folytonossága biztosított.) Az adatbázis ugyan nem tartalmazza a magyar piacon meghatározó Profession.hu adatait, ami limitációnak tekinthető. Az adatok nagyságrendjét tekintve elmondható, hogy a domináns Profession.hu (Profession.hu Kft., 2025) naponta körülbelül 14000 hirdetést tartalmaz, és ha elfogadjuk (komoly felülbecsléssel), hogy ezek mind 24 órán belül publikált, valóban új hirdetéseket jelentenek is (ami már csak a gyakori újrapiublikálások miatt szinte biztosan nem így van), akkor a vizsgált időszak alatt nagyságrendileg 20 milliós maximum hirdetésszámról beszélhetünk. Ehhez viszonyítva az 1.3 milliós, rendelkezésünkre álló minta, különösen a JobsPikr általi nemzetközi összevethetőség célját figyelembe véve, véleményünk szerint kielégítő lefedettséget jelent.

Megjegyzendő, hogy a nagy mennyiségű, nyilvános forrású adat – ezen belül is az online álláshirdetések adatai – már a nemzeti és EU szintű statisztikai hivatalok érdeklődését is felkeltették, és mint alternatív adatforrások feldolgozásra kerültek, melyről a (Daas és tsai., 2020) jelentés is beszámol. Mint ilyen, az általunk felhasznált adatforrás is ezen jelentés módszertani megközelítésébe illeszkedik, azaz az adatot nem kutatásunk céljára gyűjtötték, hanem azt „nagy adat” forrásként mi dolgoztuk fel saját következtetéseink levonására. Módszertanunk leginkább az Eurostat „Labour market demand for ICT specialists in online job advertisements” (Eurostat, 2023) nevű kísérleti statisztikai mutatójához köthető.

Az adatbázis legfontosabb változói a következők:

- **crawl_timestamp**: az adatkinyerés időpontja;
- **url**: az álláshirdetés forrásának webcíme;
- **job_title**: az állás megnevezése;
- **category**: az álláshirdetés kategóriája;
- **company_name**: a munkáltató cég neve;

- **city, state, country:** az állás földrajzi helye (város, állam, ország);
- **post_date:** az álláshirdetés közzétételének időpontja;
- **job_description:** az állásleírás szabad szöveges formában;
- **job_type:** a munkavégzés jellege (pl. teljes munkaidő, részmunkaidő);
- **salary_offered:** a kínált fizetés, amennyiben elérhető;
- **job_board:** az álláshirdetés forrása (pl. állásportál);
- **contact_email, contact_phone_number:** az álláshirdetést közzétevő kapcsolati adatai;
- **apply_url:** a jelentkezéshez vezető link;
- **logo_url:** a cég logójának elérhetősége.

Továbbá az adatbázis számos gépi úton „következtetett” változót is tartalmaz, például a *inferred_city*, *inferred_state*, *inferred_country*, *inferred_job_title*, amelyek az álláshirdetés szövegéből vagy más metainformációkból kerültek előállításra.

3.2.1 Felhasznált és kihagyott adatok

A kutatás során az alábbi változók bizonyultak relevánsnak, és kerültek felhasználásra:

- **job_description:** Az álláshirdetések teljes szöveges tartalma, amely az elemzések alapját képezte.
- **city:** A földrajzi helyszín meghatározásához használt városnév.
- **post_date:** A hirdetés publikációs időpontja, amely az idősoros elemzéseket támogatta.

Ezzel szemben az *inferred* előtaggal ellátott változókat – mint például az *inferred_city*, *inferred_job_title*, *inferred_salary* –, sőt, a kategóriát, fizetési szintet stb. leíró adatokat nem használtuk fel az elemzés során. Ezek az adatok automatikus besorolásokon alapultak, amelyek pontossága igen változóknak tűnt, és az adatelemzés során túlzott zajt eredményezhettek volna. Bár ezekben a változóknak potenciál rejlik a későbbi feldolgozás során, a jelenlegi kutatás céljaira nem tartottuk őket megbízhatónak.

A szektorális osztályozás elvégzéséhez a cégneveket előbb manuálisan normalizáltuk, összevonva a névváltozatokat és egyben kezelve például vállalatcsoportokat akkor is, ha azok eltérő leányvállalatairól van szó, majd a TEÁOR 2025 besorolási rendszert alkalmaztuk, amely egy külön programmal történt. Ez biztosította, hogy az álláshirdetések pontos iparági besorolásba kerüljenek, függetlenül a JobsPikr által nyújtott adatok pontosságától.

3.3 Szövegfeldolgozási módszertan

Elemzésünk lényege a *job_description* szöveges mező feldolgozásán alapul. A szöveg HTML elemektől – ha nem is teljes tisztasággal, de nagyrészt – megfosztott álláshirdetési szövegeket tartalmazott, melyeket elsősorban abban a tekintetben szerettünk volna osztályozni, hogy:

1. Explicit említi-e az AI vagy hozzá kapcsolódó kifejezések egyikét bármelyik nyelven.
2. Szemantikus tartalmát tekintve utal arra – akár explicit, akár implicit módon –, hogy a munkakört betöltő személy AI megoldásokkal dolgozik.

Míg az első megközelítés komplex kulcsszóalapú illeszkedésvizsgálatot jelent, és célja minden olyan hirdetés megragadása, mely valamilyen formában témaként említi az AI-t, addig utóbbi esetben olyan mély nyelvi feldolgozást kíséreltünk meg, mely tartalmát tekintve osztályozza a hirdetéseket aszerint, hogy maga az állás AI-val kapcsolatos feladatot jelent-e. E két megközelítés technikai aspektusait alább részletezzük.

3.3.1 Feldolgozási környezet

Teljes elemzésünk környezeteként egy Python 3.10.13 alapú fejlesztői környezetet választottunk, mely lehetővé tette a nagy mennyiségű bemeneti adat skálázható feldolgozását többek közt a Pandas keretrendszer felhasználásával, a nyelvi feldolgozás lépéseinek végrehajtását, valamint az ezt követő általános és idősor-specifikus elemzések elkészítését.

3.3.2 Feldolgozás reguláris kifejezésekkel

A komplex kulcsszóalapú mintaillesztés leghatékonyabb módja a (Kleene, 1951) által lefektetett alapokon nyugvó reguláris kifejezése (*Regular Expressions*, *RegEx*) használata. A RegEx kifejezések a kulcsszavak és kulcsszó-kombinációk – akár ragozott formában történő – illesztését is lehetővé teszik. Ennek szellemében megalkottunk egy reguláris kifejezést, mely az „AI”, „gépi tanulás”, „számítógépes látás”, „természetes nyelvfeldolgozás”, „adattudomány” és más, mesterséges intelligenciával kapcsolatos kifejezések különböző nyelvi változataira illeszkedik, beleértve az angol, német és magyar nyelveket, valamint egyes gyakori elírásokat is, így lehetővé téve témánk említésének kétértékű (igen/nem) osztályozását.

A reguláris kifejezések egy előre definiált mintázat alapján azonosítják és osztályozzák a szöveges adatokat. Működésük során egy adott szöveget vizsgálnak meg karakterről karakterre, és meghatározzák, hogy az tartalmaz-e az adott mintaillesztési szabályoknak megfelelő kifejezéseket. A RegEx motorok támogatják a különböző karaktercsoportokat, metakaraktereket (pl. . tetszőleges karakterre, * ismétlődésre), valamint alternatívákat és csoportokat, így lehetőség van bonyolult nyelvi szerkezetek felismerésére is. Az alkalmazás során egy álláshirdetés szövegében például megkereshetők a mesterséges intelligenciához kapcsolódó kulcsszavak és ezek variációi, lehetővé téve a

hirdetések AI-hoz való kapcsolódásának automatizált meghatározását. Ennek ellenére a szabályalapú illesztés merevsége miatt gyakran előfordulhatnak téves találatok, különösen, ha egy szó vagy kifejezés több jelentéssel is bír, vagy ha az említés csupán formai jellegű.

Az osztályozáshoz használt pontos reguláris kifejezés a függelékben található. A reguláris kifejezést skálázhatóan a Pandas keretrendszer `.apply()` megoldásán keresztül lehet hatékonyan alkalmazni a teljes adatsorra.

3.3.3 Feldolgozás LLM módszerekkel

A fenti megközelítés jelentős hátránya, hogy olyan helyzetekben is „fals pozitív” szignált ad, ahol az adott álláshirdetés csupán egyéb írásmódbeli sajátosságai miatt illeszkedik a reguláris kifejezésre (sajnos teljesen pontos fedés nem adható az ai betűkombinációk gyakori használata miatt), vagy rosszabb esetben csupán mellékesen említi az AI megoldásokat, vagy egyszerű stíuselemként (lásd alább), vagy – témánk szempontjából relevánsabban – a vállalati márkaépítés részeként. (Ez utóbbi jelenségre később kitérünk.)

Példa: AI mint félrevezető stíuselem egy álláshirdetésben

I thought about using ChatGPT to write a job ad for me and its answer was actually... Nothing. Because I decided not to use ChatGPT... Not everything can be driven by A.I., sometimes we need to speak to one another, human to human to create real connections, even in business?

We are currently looking for someone who can lend their communication and interpersonal skills to our MSP admin team to work closely with candidates and clients to build meaningful and fruitful relationships?

Nothing Artificial about this role ;) At XXX we believe working as part of great teams is the key driver to our success! Our amazing teams provide solutions and expertise that help companies position for growth, execute on strategy, and improve business agility.

Work with us and you'll be joining the market leader in Recruitment Outsourcing.

So if you are someone who: Is experienced in an administrative role or commercial support environment, preferably in staffing or recruitment...

Ennek megfelelően az álláshirdetések tartalmi alapú besorolásához mélyebb, jelentésalapú, *szemantikus feldolgozásra van szükség*. Tekintve az 1.3 millió feletti adatpont okozta skálázási kihívásokat, az emberi szemantikus kiértékelés nem megoldható, ebből következően gépi út áll csak előttünk. A problémát mint bináris osztályozási feladatot megközelítve mozgósíthatnánk a gépi tanulás teljes eszköztárát, és valószínűsíthetően elfogadható eredményeket érhetnénk el valamiféle előtanított, majd finomhangolt neurális hálós rendszerrel, melyet saját adatunkon szemantikus osztályozásra használunk. Emellett azonban az LLM modellek széles körű elterjedése lehetővé teszi hatékony és skálázható felhasználásukat szemantikus osztályozóként akár *tanítás nélküli* (*zero shot*) értelemben is, amint azt (Wang és tsai., 2023) is demonstrálta. Ez jól illeszkedik a (Ziems és tsai., 2024) által vázolt trendbe, mely az LLM-ek társadalomtudományi kutatás céljára való hatékony felhasználhatóságára mutat.

Ilyen esetben csupán megfelelő utasítást, azaz *promptot* (a fogalom eredetéhez lásd még McCann és tsai., 2018) kellene írunk, vagy esetleg *few shot* módszerrel több példát is adnunk az LLM modellnek, mely segíthet a teljesítmény maximalizálásában, amint azt (Liu és tsai., 2024) is demonstrálja. Az LLM-ek feladatmegoldási képessége nagymértékben függ attól, hogy milyen módon adjuk meg számukra az utasításokat, azaz milyen formában promptoljuk őket. A megfelelő prompt kialakítása döntő fontosságú, mivel az LLM modellek generatív jellegükből adódóan nem feltétlenül strukturált módon válaszolnak, hanem gyakran részletezhetik válaszaikat, visszakérdezhetnek, vagy magyarázatokat fűzhetnek hozzájuk. Ezért egy bináris osztályozási feladat esetén különösen fontos, hogy a prompt világosan meghatározza a kívánt kimenetet: a modellnek kizárólag az előre meghatározott két kategória (például „True” vagy „False”) egyikével kell válaszolnia, minden további kiegészítés nélkül. A strukturált válaszadás kikényszerítésére több módszer is létezik, például explicit utasítások a válaszformátumra, például: „Kérlek, kizárólag 'True' vagy 'False' értékkel válaszolj, ne adj további magyarázatot!” vagy hasonló, a modell következetességét növelő utasítások.

A bináris osztályozási feladat során az LLM promptját úgy kell beállítanunk, hogy képes legyen megkülönböztetni azokat az álláshirdetéseket, amelyek tartalmilag ténylegesen AI-hoz kapcsolódnak, azoktól, amelyekben csak felületes említés történik, vagy amelyek teljesen irrelevánsak. Ez történhet teljesen „zero-shot” módon, azaz anélkül, hogy példákat mutatnánk be neki, vagy „few-shot” tanulással, amely során néhány példát megadunk számára, hogy jobban ráhangolódjon a feladatra. A promptot finomhangolhatjuk azáltal, hogy közvetlenül belefoglalunk egyértelmű utasításokat arra vonatkozóan, hogy mit kell a modellnek keresnie a szövegben. Például megadhatjuk, hogy csak olyan hirdetéseket soroljon az AI-hoz kapcsolódó kategóriába, amelyek explicit módon megemlítik a gépi tanulást, mély tanulást, adattudományt, vagy más, mesterséges intelligenciával kapcsolatos kulcsszavakat. Alternatív megközelítésként az LLM használható szemantikus osztályozásra is, amely során nem csupán az explicit kulcsszó-egyeztetéseket vizsgálja, hanem az adott szöveg tartalmát és kontextusát is figyelembe veszi.

Az LLM modell kimenetének értelmezése szintén kulcsfontosságú, mivel az output gyakran tartalmazhat változatosságot. Például a modell egyes esetekben „True.” vagy „True!” formátumokat generálhat, ami könnyen kezelhető, azonban előfordulhat, hogy olyan válaszokat is ad, amelyek nem egyértelműen felelnek meg az elvárt bináris osztályozási kritériumoknak, például: „Ez a hirdetés valószínűleg AI-hoz kapcsolódik, tehát igen.” Ennek elkerülése érdekében a prompt megfogalmazásának egyértelműsége mellett gyakran szükség van utófeldolgozási lépésekre is, amelyek biztosítják, hogy az output megfelelő formátumban kerüljön feldolgozásra. Ez történhet például reguláris kifejezések vagy egyéb szövegelemzési módszerek segítségével, esetleg az adott szolgáltató (esetünkben OpenAI) applikációs illesztési interfészén elérhető „strukturált kimenet” opció használatával.

Az LLM-alapú osztályozók bizonytalanságait elkerülendő, választásunk ennél szigorúbb megközelítésre esett: manuális munkával előállított (a meg-

oldás sajátosságai miatt szerencsésen kis méretű) tanító és tesztadatsort állítottunk össze, majd automatikus promptoptimalizációs eljárásokkal hoztunk létre magas hatékonyságú LLM-alapú osztályozó promptot. Az optimalizált prompt segítségével elértük, hogy a modell ne csupán pontos bináris címkézést hajtson végre, hanem az előzetesen manuálisan validált tesztadatsoron is közel hibátlan teljesítményt érjen el. A következő lépésben ezt az optimalizált promptot alkalmaztuk a teljes adathalmazra, hogy automatizált módon, nagyméretű álláshirdetés-adathalmazon is azonosítani tudjuk az AI-hoz kapcsolódó hirdetéseket.

3.3.4 Promptoptimalizálás DSPy csomaggal

Az LLM-alapú feladatmegoldás ismert problémája (melyet Si és tsai., 2023 is demonstrált) hogy a modell teljesítménye nagyban ingadozhat a feladatmegoldás terén, függően az utasítások (promptok) pontos megfogalmazásától, még akkor is, ha azok emberi olvasó számára ekvivalensnek tűnhetnek. Ennek megfelelően az automatikus promptoptimalizációs eljárások széles körét fejlesztették ki (áttekintésért lásd még Li és tsai., 2024), melyek felügyeletlen vagy felügyelt tanulási feladatként kezelik az adott prompt finomhangolásának kérdését.

E célból különösen hasznos eszköz a *DSPy* nyílt forráskódú programcsomag (Khattab és tsai., 2023), mely lehetővé teszi az automatikus promptoptimalizációt.

A *DSPy* (Declarative Self-improving Python) egy nyílt forráskódú keretrendszer, amelyet a Stanford NLP csoport fejlesztett ki azzal a céllal, hogy a nyelvi modellek (LLM-ek) programozását a hagyományos promptolás helyett deklaratív, moduláris módon tegye lehetővé. Ahelyett, hogy kézzel írt promptokat alkalmaznánk, a *DSPy* lehetővé teszi a felhasználók számára, hogy a kívánt viselkedést Python kódként írják le, amelyet a rendszer automatikusan optimalizál a megadott metrikák alapján. Ez a megközelítés különösen hasznos összetett feladatok esetén, mint például a gondolatlánc jellegű feladatmegoldás (chain-of-thought) (Wei és tsai., 2023), a „retrieval Augmented Generation” (RAG) (Lewis és tsai., 2021), vagy többlépéses „ügynök rendszerek” kialakítása. A *DSPy* moduljai és optimalizálói lehetővé teszik a promptok (és akár a modellek súlyainak) automatikus finomhangolását, ezáltal csökkentve a kézi beavatkozás szükségességét, és növelve a rendszer megbízhatóságát és teljesítményét.

Munkánk során e csomag felhasználásával, felügyelt tanulási feladatként megfogalmazva egy 150 AI-hoz nem kapcsolódó és 50 AI-hoz kapcsolódó álláshirdetésből álló tréning adatszettet, és egy 10 példából álló teszt adatszettet állítottunk össze, ügyelve arra, hogy a fentiekhez hasonló, félrevezető hirdetések, és hasonló, komoly kihívást jelentő adatpontok is szerepeljenek a tesztben. A *DSPy* keretrendszer segítségével a *COPRO* (*Cooperative Prompt Optimization*) optimalizációs eljárással 100%-os teszt F1 metrikájú, azaz a feladatot a tesztadaton hibátlanul megoldó osztályozót hoztunk létre.

Megoldásunk létrehozásához a *COPRO* optimalizációs eljárást használtuk.

A COPRO (Cooperative Prompt Optimization) módszer a DSPy keretrendszerben elérhető optimalizációs algoritmus, amely a promptok iteratív finomhangolását végzi együttműködő tanulási stratégiák alkalmazásával. A módszer kulcsfontosságú eleme, hogy egy együttes optimalizálási folyamat során nem csupán a promptokat módosítja, hanem az azokat értékelő modellt (illetve esetünkben promptot) is adaptívan hangolja, biztosítva ezzel a célfeladatra való maximális illeszkedést. A COPRO többkörös iteráció során az adott adathalmazon elért teljesítmény visszacsatolásából tanul, és az optimalizált promptokat folyamatosan finomítja. Ezzel lehetővé teszi, hogy az így kapott promptok az alapmodell képességeit a lehető legjobban kihasználják, és minimálisra csökkentsék az olyan hibázási lehetőségeket, mint az adott kontextusban irreleváns vagy félrevezető osztályozás.

Az eljárás során először az eredeti promptból kiindulva több (alapértelmezetten 10) alternatív utasítás generálása történik, melyekhez a rendszer automatikusan hozzáilleszti az utolsó kimeneti mezőhöz tartozó ún. *prefix* részt is. Ezután minden alternatív promptot egyenként értékelünk a megadott metrikával, a teljes program újrafuttatásával. Az így szerzett visszacsatolás alapján a legjobban teljesítő utasításokat kiválasztjuk, és ezekből példákat generálva (few-shot stílusban) újabb utasítás-jelölteket hozunk létre. Ez a folyamat több (pl. 3) mélységi iteráción keresztül ismétlődik. A kiválasztás során nemcsak az abszolút legjobb, hanem a legfrissebb promptgenerálási körök átlageredményeit is figyelembe vesszük. A kód lehetőséget nyújt statisztikai mutatók (minimum, maximum, átlag, szórás) követésére is az optimalizációs folyamat során.

Az eljárás eredményeként optimalizált promptkifejezés a függelékben található.

Ezt követően a promptot az *OpenAI GPT-4o-mini* modelljét az OpenAI Platform API-ján keresztül (OpenAI, 2025) használtuk fel, és (a Tier 4 limitációk okán) közelítőleg 200 órás futásidővel lefuttattuk a teljes 1.3 milliós adatállomány bináris szemantikus osztályozására.

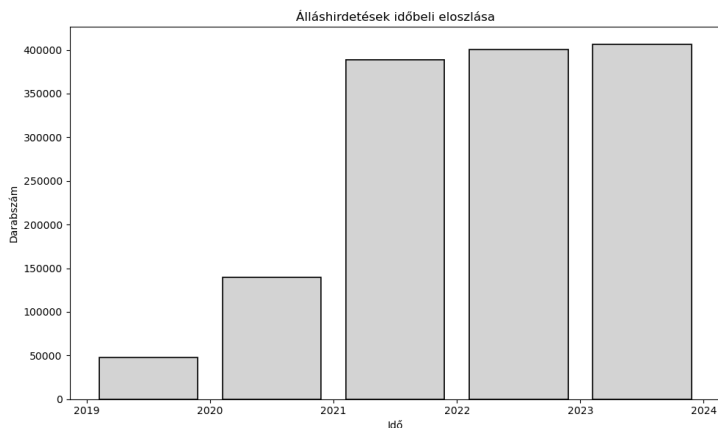
Alábbi elemzéseinkben a kulcsszóalapú megoldásra RegEx néven, a fenti LLM-alapú megoldásra DSPy néven fogunk hivatkozni.

4 Elemzés

4.1 Feltáró adatelemzés

Az adatok elemzésének első lépéseként feltáró adatelemzést végeztünk, vizsgálандó az álláshirdetések térbeli és időbeli eloszlását.

Az adatbázis összesen 1382827 álláshirdetést tartalmazott, a publikálás dátuma szerint az alábbi eloszlásban (lásd 1. ábra).

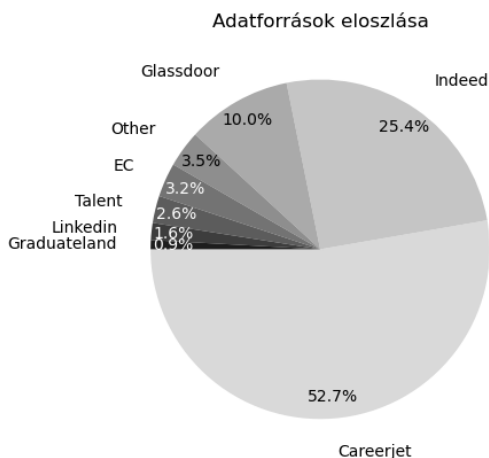


1. ábra. Az álláshirdetések publikálásának időbeli eloszlása

Mint látható, az adatállomány számossága a 2019-es és a 2020-as évben nem teljesen kiegyenlített a többi évhez képest (ami adott esetben a trend-előrejelzések során okozhat problémákat), ám így is igen nagy számosságról van szó, a 2021-es évtől kezdődően kifejezetten nagy mennyiségű adat áll rendelkezésre.

Az adatok területi eloszlása jelentős egyensúlytalanságot mutat Budapest abszolút dominanciájával. (Az adat területi alapú vizualizációja épp ezért szinte lehetetlen, Budapestre vonatkozó hirdetések száma még a nagy megyeszékhelyekének is sokszorosa, teljesen torz vizualizációhoz vezetve.)

Ez jól mutatja az adat abbéli limitációját, hogy csak online forrásokból, azon belül is kifejezetten külföldi tulajdonú online állásportálok dominálnak. Egyrészt fontos kiemelni, hogy akár a Profession.hu és egyéb nagy magyar portálok is hiányoznak, másrészt ez az orientáció természetesen eltolhatja az adatot a nemzetközibb orientációjú munkák felé. Szerencsére tüzetesebb elemzés során azért bőséggel lehetett találni alacsonyabb végzettséget igénylő munkákat éppúgy, ahogyan Budapesten kívülieket is (lásd 2. ábra).



2. ábra. Az álláshirdetések adatforrásainak eloszlása

A leggyakoribb 15 város elemzése jól mutatja Budapest dominanciáját, ám azért egy általánosan jó lefedést is mutat az ország egészét illetően (lásd 3. ábra).



3. ábra. Top15 település az álláshirdetések gyakorisága szerint

4.2 Általános trendek

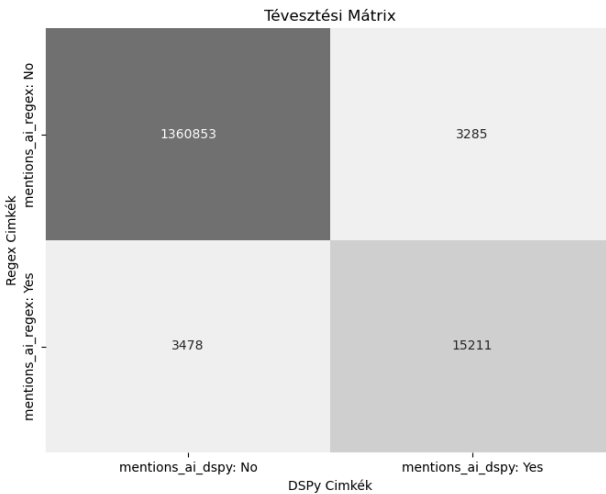
Elemzésünk legfontosabb fókuszja az AI-jal kapcsolatos álláshirdetések szektorális eloszlása, és mind általános, mind pedig egyes kiemelt cégeket érintő időbeli mozgásának elemzése. Ahhoz azonban, hogy érdemi képet tudjunk

alkotni ezekről, érdemes elsőként összevetnünk a két kidolgozott módszer, tehát a RegEx és a DSPy alapú osztályozás eredményeit, illetve ezek eltéréseit, ami fényt vethet a módszerek egyedi előnyeire is.

4.2.1 RegEx és DSPy módszerek közti eltérések

Mint azt korábban említettük, a DSPy alapú osztályozó a kiválasztott teszt-adatokon igen jól teljesített, és a teljes adatsoron történő alkalmazása után történt szűrőpróbaszerű vizsgálatokon is kifejezetten jól szerepelt, így – nem meglepő módon – vizsgálataink során erősebben fogunk rá támaszkodni.

A RegEx és DSPy alapú megoldások közti „tévesztési” (jobban mondva talán „egyet nem értési”) mátrix a 4. ábra szerint alakul.



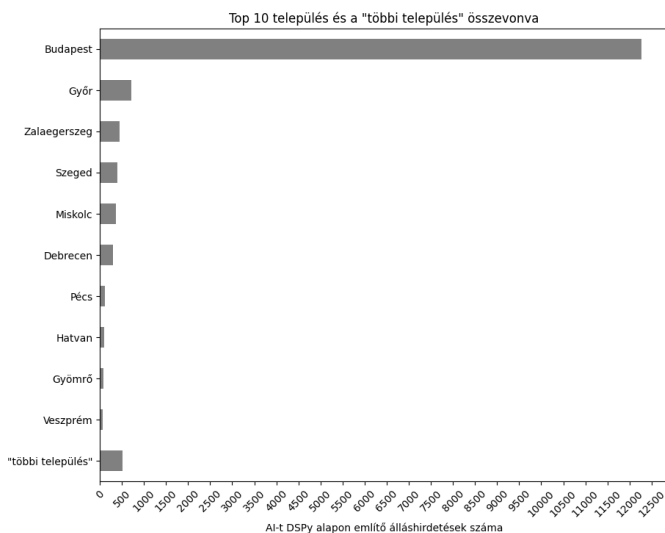
4. ábra. RegEx és DSPy megoldások közötti azonosság/eltérés tévesztési mátrix formában

Mint látható, 3000 feletti mindazon bejegyzések száma, melyeket a RegEx megoldás egyszerűen nem azonosított, például mivel csak a szöveg jelentéséből, nem pedig kulcsszó szerinti egyezésből adódóan van köztük az AI-hoz, vagy csupán felszíni (néha kifejezetten csak véletlenszerű) szóegyezés miatt pozitív jelet adott rájuk. Mint ilyen, úgy érezhetnénk, hogy a RegEx megoldás eredményeit nem érdemes használni, azonban az alábbiakban később tárgyalt módon rámutathatnak egy speciális esetre, melyben a cégek kifejezetten tudatosan használják az AI-hoz köthető kifejezéseket hirdetéseikben, még akkor is, ha maga a hirdetett állás nem kapcsolódik a témához.

Összességében az adatsorban (DSPy alapon) *18496 AI-hoz köthető hirdetés található, ami a teljes hirdetésállomány elenyésző, 1.34%-os arányát jelenti. Ennyiből is megállapítható, hogy a közvetlen AI technológiákkal való munka a magyar álláshirdetések igen kis százalékában van jelen.*

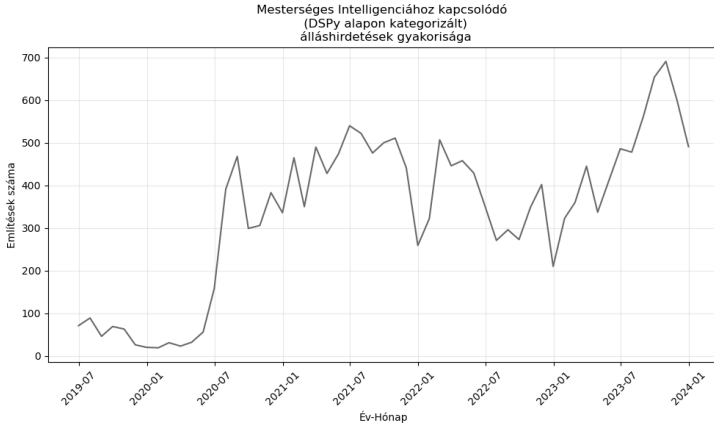
4.2.2 Általános és időbeli eloszlás

Elsőként tekintetbe véve a AI-hoz köthető álláshirdetések területi eloszlását látható, hogy Budapest általános dominanciája még a korábbiaknál is erősebb. Ezt természetesen nagyban befolyásolja a tény, hogy maga az alapadat is erősen Budapest-fókuszú, bár a következtetés, hogy AI használat terén a fővárosi vállalkozások járnak előrébb, nem tűnik légből kapottnak, annál is inkább, mert a hirdetések általános gyakoriságához képest az AI-hoz köthetőek arányában Budapest még drámaibban dominál (lásd 5. ábra).



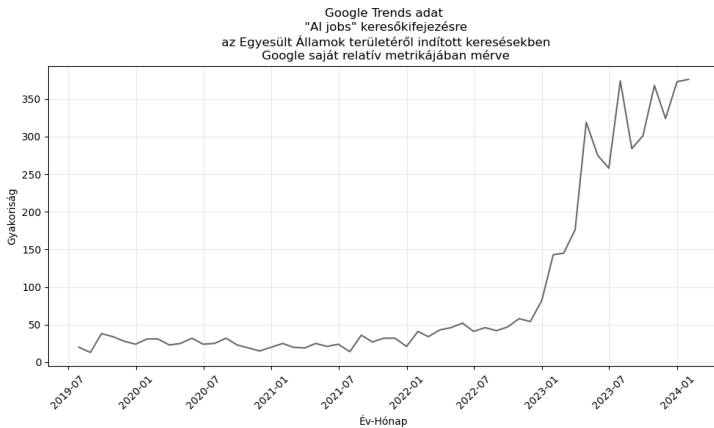
5. ábra. Mesterséges Intelligenciához kapcsolható (DSPy alapon kategorizált) álláshirdetések területi eloszlása

Az összes DSPy alapon AI-hoz köthető álláshirdetés időbeli eloszlásának vizsgálata (lásd 6. ábra) fontos eredményekre hívja fel a figyelmet: Csupán a darabszámot figyelembe véve két jelentős „hullám” figyelhető meg az idősorban: 2020-ban és 2023-ban. Míg az első hullám valószínűleg a „klasszikus”, azaz nem „megalapozó modell” (vagy elterjedtebb nevén: „generatív AI”) megoldásokhoz köthető, és adott esetben magyarázható az (alább kifejtett) autópári szektor fellendülésével, addig a 2023-as hullám igen valószínűleg az OpenAI ChatGPT modelljének 2022 novemberi bejelentésével és az ezt követő „érdeklődéscunami” jelenségével magyarázható, mely a közbeszédet nagyban áthatotta. Ennek jó illusztrációja a Google keresőkiejezések használati gyakorisága mind a tengerentúlon (lásd 7. ábra), mind pedig a magyar keresésekben (lásd 8. ábra) (Forrás: Google, 2025.)

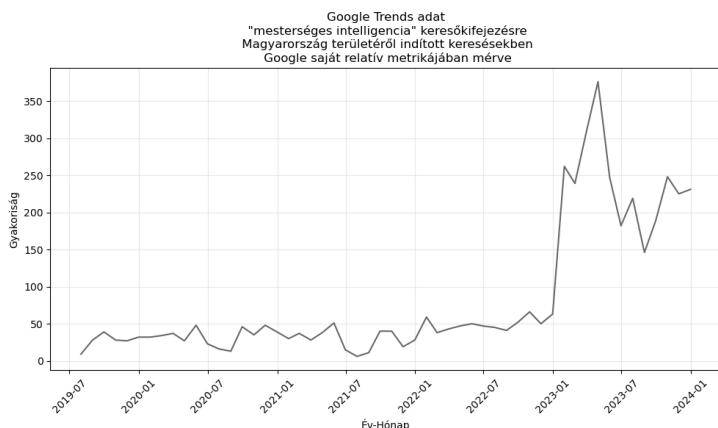


6. ábra. Mesterséges Intelligenciához kapcsolódó (DSPy alapon kategorizált) álláshirdetések gyakorisága

Megjegyzendő azonban egy fontos különbség, melyet a specifikus keresőkifejezés megválasztásánál mi is figyelembe vettünk: csupán az AI témára általában keresve a tengerentúli érdeklődés hatalmas, de ChatGPT-hez kapcsolható felívelése egybeesik az „AI jobs” kifejezések igen aktív felhasználásával. A Magyarországra vonatkozó ábrán a „mesterséges intelligencia” témájáról általában véve vett kereséseket láthatjuk, mert a „mesterséges intelligencia munka” és hasonló keresőkifejezések egészen elenyésző, értelmezhetetlenül kevés adatot eredményeznek.

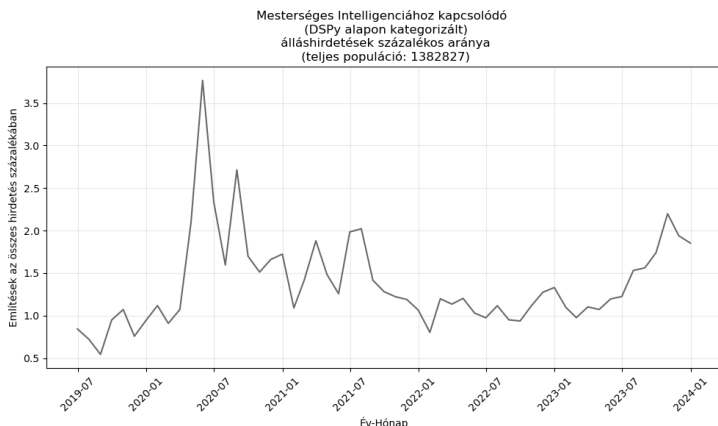


7. ábra. Google Trends adat „AI jobs” keresőkifejezésre az Egyesült Államok területéről indított keresésekben Google saját relatív metrikájában mérve



8. ábra. Google Trends adat „mesterséges intelligencia” keresőkifejezésre Magyarország területéről indított keresésekben Google saját relatív metrikájában mérve

Ez a jelenség arra utalhat, hogy míg a tengerentúlon a ChatGPT modell (és evvel az egész „generatív AI” trend) érkezését a munka világra közvetlenül vonatkozóan, mintegy „praktikus alkalmazására fókuszálva” vette figyelembe a közvélemény, addig a magyar közbeszéd sokkal inkább mint a munkától távolabbi „kuriózumként” tekintett e modellekre, és a magyar munka világában továbbra is a „klasszikus”, egyedi, célorientált AI fogalma dominál. Ennek hatása megfigyelhető abban is, ha a magyar AI álláshirdetéseket nem abszolút számosságukban, hanem relatíve, az összes hirdetés százalékában vesszük szemügyre (lásd 9. ábra). Ebben a nézetben a 2023-as hullám bár észrevehető, ám magnitúdójában bőséggel elmarad, ez utalhat arra, hogy – legalábbis az elemzés adatsorunk által limitált pillanatában – a „generatív AI” hulláma 2024 Magyarországra még „nem ért ide”.



9. ábra. Mesterséges Intelligenciához kapcsolódó (DSPy alapon kategorizált) álláshirdetések százalékos aránya

4.3 Hol vannak a legújabb technológiák?

Kérdés, hogy az AI technológiák legújabb inkarnációi, a közvélemény és a média által – nem túl precízen – „Generatív AI”-nak nevezett megoldások (melyek esetében helyesebb lenne általánosan a Bommasani és tsai., 2022 értelmében vett „megalapozó modellekről”, vagy specifikusan leggyakrabban „nagy nyelv-modellekről” beszélni) milyen mértékig jelennek meg a hirdetésekben. Tüzetesebb keresés után megállapítható, hogy a „ChatGPT” mint specifikus megoldást és az LLM technológiákat általános értelemben érdemben 128 hirdetés, „Generatív AI”-t 106 hirdetés említi összesen, azaz elmondható, hogy e technológia munkaerőpiaci penetrációja még igen gyenge.

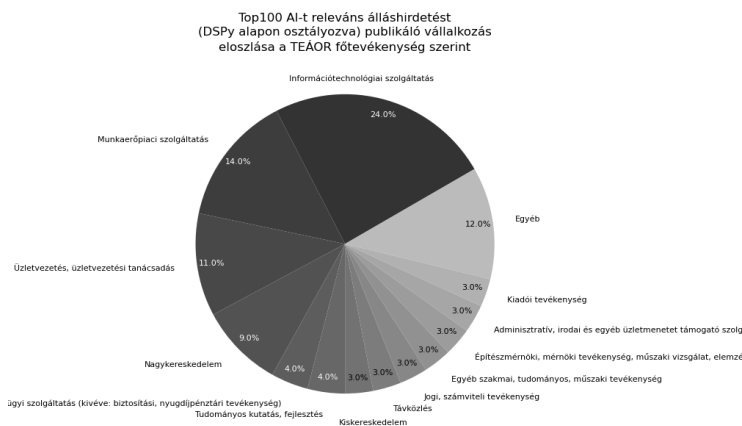
Itt megjegyzendő, hogy amennyiben azonban ezeket általános célú technológiáknak tekintjük, úgy könnyen lehet, hogy számos hirdetésfeladó gondolhatja „említésre nem is érdemesnek” az állás szempontjából a ChatGPT vagy egyéb LLM technológiákat. E lehetőséget nem tudjuk kizárni, ám megemlítendő, hogy az Excel kifejezést (RegEx alapon, csak önálló, nagybetűs szóként) tartalmazó hirdetések száma 168348, ami mégiscsak az összes hirdetés 12.17%-a, annak ellenére, hogy gyakran érezhetjük e technológiát már-már a „kötelező irodai minimum” részének.

Ebből a szempontból érdemes lesz további kutatásokat végezni a legújabb AI technológiák beépülése szempontjából, illetve megállapíthatjuk, hogy a fenti, 100-as nagyságrendű említés erősen arra utal, hogy e technológiák még egyszerűen nem értek el a magyar vállalatok hétköznapijaiba.

4.4 Szektorális eloszlás

Különösen érdekes vizsgálati szempont az AI témával kapcsolatosan leggyakrabban hirdető szervezetek szektorális eloszlása. Annak érdekében, hogy ez vizsgálható legyen, manuális adatmegfeleltetéssel feldolgoztuk az adatban szereplő vállalkozásnevek zömét (minden olyan vállalkozást, mely 5-nél többször adott fel AI-hoz kapcsolható hirdetést), összevonva neveik összes ismert írásmódját, ügyelve arra, hogy egyes cégcsoportok esetén azok alvállalkozásai is egy név alá kerüljenek. Ezt követően a nyilvánosan elérhető cégalapokból kikerestük az AI témájú hirdetésfeladás (DSPy és RegEx alapon is) leggyakrabban 100-100 vállalkozás adatait, főtevékenységük TEÁOR’25 szerinti besorolását, majd ennek első két számjegyét alapul véve készítettünk szektorális elemzést (lásd 10. ábra). Az elemzés során a teljes TEÁOR kódot oly módon dolgoztuk fel, hogy első két számjegyét emeletük ki (azaz a pontot követő alrészeket levágtuk, és az első két számjegyet vettük figyelembe). Emellett a jobb átláthatóság érdekében a 62-es és 63-as alcsoportot egységes informatikai csoporttá vontuk össze, valamint a 45-ös alcsoportot (jármű-kiskereskedelem) a 47-es alcsoportba (kiskereskedelem) olvasztottuk. Egyéb szerkesztés nem történt, így az alábbi elemzés a főbb gazdasági ágazatok szerint mutatja az AI hirdetések gyakoriságát. Az ágazatok időbeli változásait nem elemeztük, mivel nem éreztük biztosítva az állandó ágazati lefedettséget az adatok időbelisége tekintetében. Elemzésünket újabb adatszett gyűjtésével érdemes lehet majd későbbi időpontban is megismételni (ezért is választottuk

a TEÁOR 2025 szerinti besorolást, annak ellenére, hogy az adat korábbról származik), így későbbiekben megállapítható lehet az időbeli szektorális dinamika, és akár a COVID időszak hatása is láthatóvá válhat. Ezt azonban későbbi adatgyűjtéssel összekötött új elemzés körébe soroljuk.



10. ábra. Top100 AI hirdetést (DSPy alapon) feladó vállalkozás szektorális eloszlása

A szektorális eloszlás érdekes tanulságokat rejt, bár torzításokat is tartalmaz: először is elég szembetűnő, bár nem meglepő, hogy az IT terület dominál, azaz a top 100-ban való 24%-os részesedésével messze a legtöbb AI-hoz kapcsolódó hirdetés az információs technológiai szektor szereplőitől származik. Ez egyrészt várható eredménynek tekinthető, mivel e szektor felel nem kis részben a többi szektor AI használatának elősegítéséért is – például beszállítóként –, másrészt természetesen „közelebb esik a témához”, azonban valamilyen értelemben „early adopter”-nek (Rogers, 2003) is tekinthető. Ebből adódik, hogy későbbi fókusz elemzések során az AI adoptációja tekintetében érdemes erre a szektorra koncentrálni.

A második helyen álló „munkaerőpiaci szolgáltatás” szektor viszont igen csalóka: elemzésünk szerint csupán azért került ilyen előkelő helyre, mivel az összes olyan hirdetést, melyet munkaerő közvetítéssel, „fejvadászattal” foglalkozó vállalkozás ad fel, és nem kívánja nyilvánosan megadni megbízója nevét, az adatbázis kénytelen volt magához a munkaközvetítő vállalkozáshoz sorolni. Így a szektor által képviselt 14% tulajdonképp egyfajta „ismeretlen” kategóriát kellene képviseljen.

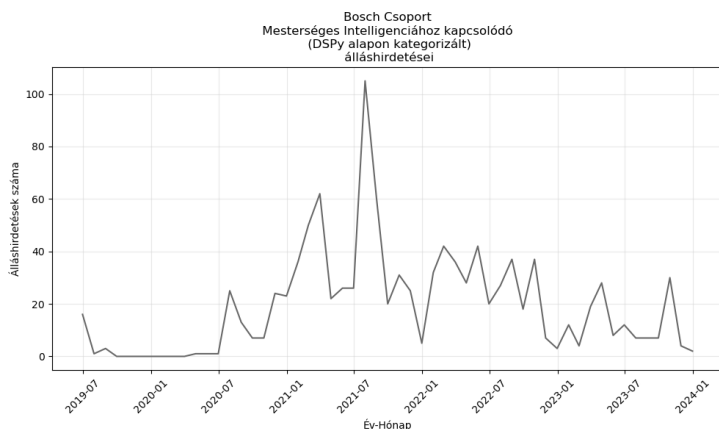
Harmadik helyen egyértelműen a tanácsadási szektor áll, nem csupán amiatt, mert maga is progresszív, „technológiaközel” szektornak tekinthető, hanem ismételten azért, mert valószínűsíthetően a tanácsadó cégek maguk azért építenek AI kapacitást, hogy szolgáltatásaik részeként építsenek megoldásokat ügyfeleik részére. Véleményünk szerint ez a szektor képviseli az IT mellett legerősebben az *AI beszállítói réteget*.

További említésre méltó területeknek tekinthetőek a banki, nagykereskedelmi (főként energiakereskedelmi) és kutatási (főként gyógyszerkutatási) szektorok, egyrészt mert tőkeerejük miatt talán a leginkább képesek előre-

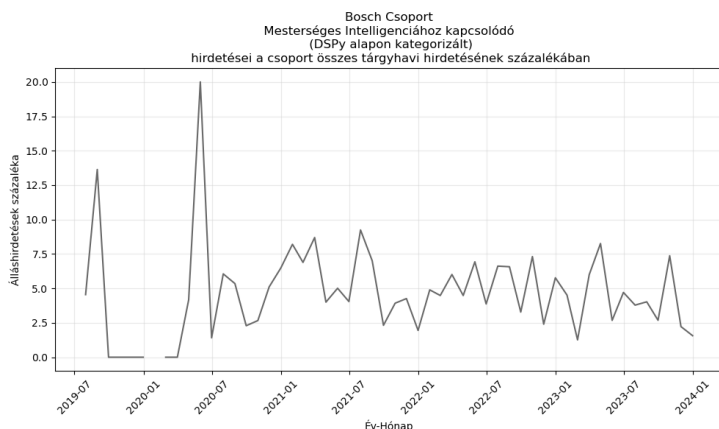
4.5.1 Egyedi elemzések, időbeli dinamika

„Egy hullám” stratégia Tovább erősíti hipotézisünket az autóipar meghatározó szerepéről, ha közelebbről megvizsgáljuk a Bosch-csoport AI hírdetéseinek időbeli eloszlását mind abszolút értékben, mind százalékosan tekintve (lásd 12., 13. ábra): míg abszolút értékben a 2021-es év tekinthető kiemelkedőnek, az is szembeűnő, hogy a 2020-as év egy szakaszában az összes, Bosch-csoporthoz tartozó hirdetés közel 20%-a AI relevanciájú volt, mely időben egybeesik a cégcsoport önvezetést célzó erőfeszítéseinek felskálázásával, és bősséggel megelőzi a ChatGPT 2022 novemberi berobbanását. Bár az önvezetés témája korábbról is kiemelt szerepet játszott a Bosch magyarországi tevékenységében, így egzakt időzítés nem könnyen indokolható, de a téma erejét jól illusztrálja például egy sajtóbejelentés (Robert Bosch Kft., 2021) 2021 elejéről.

Ez a „hullámszerű” igény az AI-ra, mely rövid időn belül szaturálódik, egyfajta fix keretű kompetenciaépítést sugall.

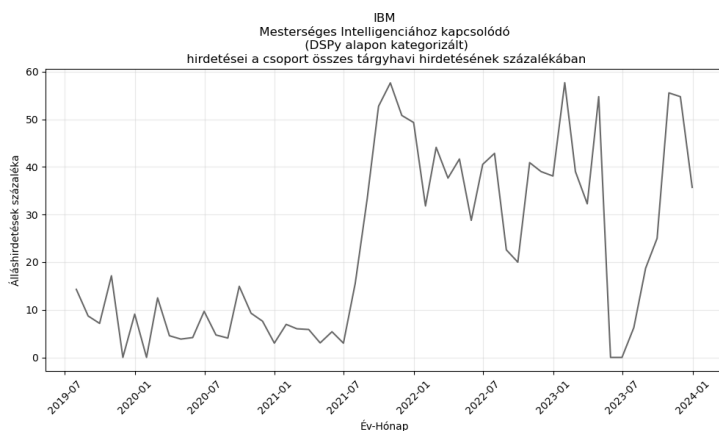


12. ábra. Bosch-csoport AI-hoz köthető hírdetéseinek darabszáma



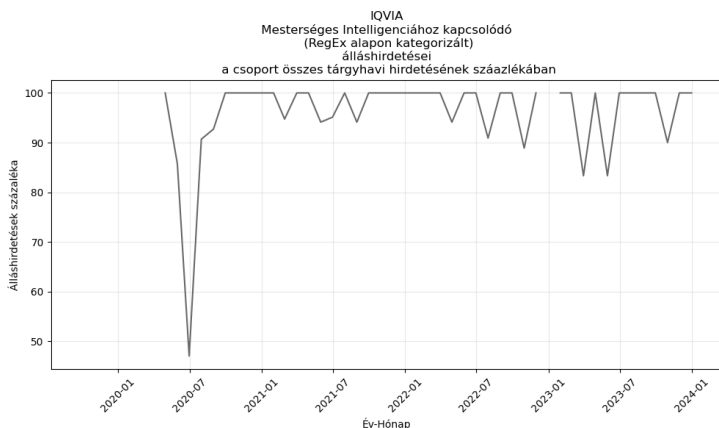
13. ábra. Bosch-csoport AI-hoz köthető hírdetéseinek százalékos aránya

„Átállás” stratégia Érdekes azonban ezt összevetnünk például az IBM stratégiájával, mely különösen a százalékosan normalizált adatsorból szembeűnő (lásd 14. ábra). Az IBM stratégiája, úgy tűnik, hogy egy kezdetben alacsonyabb kiindulási szint után az AI-képes munkaerő tartósan magas százalékon való tartása, azaz egyfajta folyamatos felvétel (vagy legalábbis folyamatos tehetségkeresés), ami arra utalhat, hogy a vállalat elégséges piaci keresletet érzel 2021 közepétől kezdődően ahhoz, hogy stratégiai erőforrásként felépítse és aktuálisan bővítse / frissen tartsa AI-képes munkaerő-kapacitását. Tulajdonképpen azt is mondhatjuk, hogy a vállalat egy váltópont után „átállt” egy bizonyos, többé-kevésbé állandó szintű felvételre (vagy keresésre) a témában.



14. ábra. IBM csoport AI-hoz köthető hirdetéseinek százalékos aránya

„Company branding” stratégia Ha azonban a vizsgálat körét a szigorúan vett (DSPy alapon megállapított) top vállalatok halmazán túl bővítjük, felfigyelhetünk egy érdekes esetre: az IQVIA cég ugyan a DSPy alapon nem tartozik a lehető legtöbb keresők közé (bár messze nem elhanyagolható, „középmezőnyt” képvisel), igen érdekes megfigyelni, hogy RegEx alapon, tehát kulcsszómegjelenés tekintetében egyenesen a legdominánsabb cég (top 1). RegEx-alapú időbeli, százalékos adata (lásd 15. ábra) pedig rövid felfutás után maximális, 100%-os szaturációt mutat.



15. ábra. IQVIA csoport RegEx alapon AI-t említő hirdetéseinek százalékos aránya

Értelmezve a látottakat: bár tartalmát tekintve messze nem arról van szó, hogy az IQVIA ezek után csak AI szakembereket keres a munkaerőpiacon (DSPy alapon nincs a leggyakoribbak közt), ám a vállalat minden hirdetésben konzisztensen megemlíti az AI témáját, a vállalati márkáépítés részeként, azaz úgy érzi, hogy a potenciális munkavállalók felé ez pozitív, vonzó üzenet, a cégimázs szerves része.

A fenti egyedi példák jól szemléltetik, hogy különféle eltérő stratégia létezik a vállalatok számára az AI képességek munkaerő körében történő beemelésére.

4.6 Csoportalkotási kísérletek

A fenti manuális „stratégiakeresés” kiegészítéseként alapos kvantitatív alapú keresést is folytattunk, hogy fedezhető-e fel további informatív és interpretálható csoportosítás a DSPy alapon megállapított AI hirdetések cégenkénti százalékosan normalizált idősoros adataiban. Ehhez elsőként a „nyers” havi százalékos idősorokon hajtottunk végre DBSCAN algoritmussal (Ester és tsai., 1996) klaszterezést, mely során az algoritmus optimális ε hiperparaméterének megállapítását (Satopää és tsai., 2011) „Kneedle” algoritmusára bíztuk, majd megismételtük ezen elemzést a „Dynamic Time Warping” (Sakoe & Chiba, 1978) előfeldolgozott páronkénti hasonlósági mátrixon, ezt követően pedig a „Matrix Profile” (Yeh és tsai., 2016) megoldás által előfeldolgozott adaton, valamint ennek negyedéves simított változatán. A fenti felismeréseken túl azonban számottevő új stratégiát nem sikerült azonosítanunk.

4.7 Munkakörök elemzése

Ahhoz hogy részletesebben feltárhassuk a hirdetések mögött rejlő konkrét szerepeket, elemeztük a szövegekben előforduló, illetve azokból kikövetkeztethető konkrét szerepkörök eloszlását. Ehhez ismét a DSPy módszer által azonosított AI-hoz köthető hirdetések körét (összesen 18496 hirdetést) vettük

megjelent a tendencia, melyben több vállalat fektet be mesterséges intelligencia vezérelt termékek fejlesztésébe és piacra vitelébe, vagy belső termékfejlesztésbe.

Érdekes továbbá megjegyezni, hogy a mesterséges intelligencia nem kizárólag informatikai területeken jelenik meg. Olyan, hagyományosan nem IT-központú pozíciók is feltűnnek, mint a *tanácsadó* vagy akár a *könyvelő*. Ez arra utalhat, hogy a mesterséges intelligencia beépülése a szélesebb munkavégzésbe – ha igen lassan is –, de megkezdődhetett.

Végül az álláshirdetésekből észlelt *gyakornoki szintek* jelenléte azt jelzi, hogy a vállalatok aktívan törekszenek a mesterséges intelligenciával kapcsolatos tehetséggondozásra és az utánpótlás biztosítására. Ez hosszú távon hozzájárulhat az AI ökoszisztéma folyamatos fejlődéséhez, és az iparág innovációs potenciáljának fenntartásához.

A pozíciók hierarchiájának szisztematikusabb kialakítása és esetleges mélységi elemzése további kutatás tárgya lehet.

5 Következtetések

Fenti eredményeinket összegezve elmondhatjuk: a kutatás során megvizsgáltuk a mesterséges intelligencia (AI) munkaerőpiaci penetrációját Magyarországon, nyilvános online álláshirdetések szövegelemzésén keresztül. A kapott eredmények alapján az alábbi fő következtetéseket vonhatjuk le:

- *Az AI-hoz kapcsolódó álláshirdetések aránya alacsony:* A teljes adatbázisban az AI-hoz köthető hirdetések aránya mindössze 1.34% volt, ami arra utal, hogy a mesterséges intelligenciával kapcsolatos munkahelyek aránya jelenleg még alacsony Magyarországon.
- *Területi koncentráció:* Az AI-hoz kapcsolódó álláshirdetések rendkívül erős fővárosi koncentrációt mutatnak. Budapest dominanciája a teljes adatbázisban is jelentős, de az AI-hoz kapcsolódó hirdetések esetében még inkább kiemelkedő.
- *Időbeli trendek:* Két fő hullámot azonosítottunk az AI-hoz kapcsolódó álláshirdetések időbeli eloszlásában:
 - Egy 2020-2021-es csúcspont, amely feltehetően az autóipari fejlesztésekhez köthető.
 - Egy 2023-as kisebb megugrás, amely valószínűleg az OpenAI ChatGPT bevezetésének hatását tükrözi, de nem mutat olyan drasztikus emelkedést, mint a nemzetközi trendek.
- *Szektorális eloszlás:* Az AI-hoz kapcsolódó hirdetések főleg az IT szektorhoz (24%), tanácsadáshoz (11%), valamint a banki és nagykereskedelmi szektorhoz kapcsolódnak.

- *Domináns vállalatok:* Az AI-hoz kapcsolódó munkahelyeket leginkább hirdető vállalatok közül a Bosch, a Continental és az AIMotive vezető szerepet tölt be, ami megerősíti az autóiipari fejlesztések dominanciáját ezen a területen.
- *Egyedi stratégiák:* A cégek különböző stratégiákat követnek az AI alkalmazásával kapcsolatban:
 - Egyes vállalatok „hullámszerű” AI felvételi stratégiát követnek.
 - Vannak, akik inkább „fokozatos átállást” mutatnak, kezdeti fel-futást követően az AI-hoz kapcsolódó hirdetéseik aránya folyamatosan magas szinten maradt.
 - Végül van eset, ami egy „company branding” stratégiát támaszt alá, ahol minden álláshirdetésben említik az AI-t, függetlenül attól, hogy maga a munkakör közvetlenül AI-hoz kapcsolódik-e.
- *Generatív AI beágyazottsága:* A legújabb AI technológiák (például Chat-GPT) jelenléte az álláshirdetésekből elenyésző (mindössze 128 explicit említés), ami azt sugallja, hogy ezek a megoldások még nem társultak a magyarországi munkaerőpiaci gyakorlatokhoz.
- *Módszertani validálás:* Az LLM (DSPy) alapon optimalizált szemantikus osztályozás pontosabbnak bizonyult, mint az egyszerű kulcsszó-alapú RegEx megoldás. A RegEx megoldás felhívhatja a figyelmet az AI említések márkaépítés célú használatára, de sok „fals pozitív” eredményt is generál.

Az eredmények azt mutatják, hogy az AI jelenléte a magyar munkaerőpiacon szektorálisan koncentrált, területenként erősen Budapest-központú, és az AI bevezetési stratégiája vállalatonként eltérő. Az AI technológiák egyre népszerűbbek, de elterjedtségük még nem éri el a tömeges alkalmazás szintjét, különösen az új, generatív AI megoldások tekintetében.

6 Limitációk

A kutatásunknak több jelentős limitációja van, amelyeket az alábbiakban részletezünk.

Az egyik legfontosabb korlát az AI megoldások egyre általánosabb jellege. Korábban az AI megoldások célorientáltak voltak, míg az utóbbi években megjelentek a nagy nyelvi modellek (LLM-ek), amelyek univerzális asszisztensként és problémamegoldóként működnek. Ez a változás hatással van az AI-expozíció mérési módszereire is.

Munkájukban Felten és társai (E. W. Felten és tsai., 2018, 2019, 2021), valamint az ezt követő frissített, bővített munkák, például (E. W. Felten és tsai., 2023), az AI és a foglalkozások kapcsolatát az O*NET taxonómián alapuló megközelítéssel vizsgálják. Ehhez az EFF AI fejlődési méréseit és Amazon Mechanical Turkon végzett adatrögzítési folyamatokat használják, amely

egy „AI occupational exposure” (AIOE) pontszámot eredményez. (Hasonló kutatási programot követ még Pizzinelli és tsai., 2023 és Gmyrek és tsai., 2023 is.)

A jelenlegi kutatásunk az álláshirdetéseket nem univerzális LLM „asszisztensek” feltételezve vizsgálta, hanem kifejezetten olyan módon, hogy a hirdetés szövegében az AI-t explicit módon említik (Regex), vagy az AI célirányos rendszerek használatát impliciten (szemantikájukban) feltételezik (DSPy). Ahogy az AI technológiák fejlődnek, egyre általánosabbá válnak, egy ponton olyan természetesen beépülhetnek a mindennapi munkafolyamatainkba, mint napjainkban az „Office” vagy egyéb „irodai programcsomagok”: széles körben használt, de ritkán említett eszköz, melyet egyszerűen nem említenek egy álláshirdetésben, hacsak nem igen alacsony képzettséget vagy junior szintet feltételező hirdetésről van szó (ott sem szó szerint említve). Ez azt jelenti, hogy a jövőben az AI-expozíció elemzése értelmetlenné válhat az explicit említések alapján, és valamilyen módon vissza kell nyúlnunk az egyes állások leírásának közvetett elemzéséhez, feltételezésekkel élve, hogy milyen aktuális technológia használatát implikálják.

Ennek érdekében a Felten-csoport fent említett kutatási módszerét LLM-alapú feldolgozással (akár DSPy segítségével) felskálázva alkalmazni tudjuk a hirdetési korpusz elemzésére. Ez erős jövőbeni kutatási irányt jelenthet, igen hasonló módszert követve, mint (Eloundou és tsai., 2023), ám magával vonja a veszélyt, hogy a technológia képességeit és azok leképeződését szubjektív módon állapítjuk meg, így előítéleteinket az LLM utasításaiba is „beégetjük”. Minden fenntartás ellenére az LLM-alapú feldolgozás erőforrásigényének jelentős csökkenése miatt ez továbbra is ígéretes kutatási irány lehet.

Egy másik fejlesztési lehetőséget jelenthet eredményeink elmélyítésére az eredeti adatsorban tartalmazott, főként automatikusan generált változók potenciálisan zajszűrés utáni feldolgozása, melyek további, a munkakörök szenioritását, hozzájuk tartozó kompenzáció esetleges mértékét tehetik elemezhetővé.

További gyümölcsöző kutatási irányt jelenthet a szövegek kifejezés szintű szemantikai elemzése abból a szempontból, hogy feltárható legyen, milyen fogalmak vagy üzleti területek állnak a legközelebb az AI témához a hirdetések szövegében, potenciálisan fényt vetve a leginkább érintett üzleti területekre/folyamatokra.

7 Szakpolitikai javaslatok

Az elemzés eredményei alapján alátámasztható a nézet, hogy a mesterséges intelligencia (AI) munkaerőpiaci adoptációja Magyarországon alacsony, és főként az informatikai szektorhoz kötődik. Több jel utal arra, hogy a vállalkozások nem tekintik általános célú technológiának az AI-t, különösen a generatív AI megoldásokat, melyek alkalmazása így a mindennapi üzleti folyamatokban még nem elterjedt. Mindezek fényében néhány jól körülhatárolható szakpolitikai intézkedés javasolható az AI magyarországi munkaerőpiacon történő erő-

teljesebb adoptációjának előmozdítására:

- *Tájékoztatási és képzési programok.* Széles körű tájékoztatási kampányok indítása az AI üzleti és ipari alkalmazásainak bemutatására. Az iparkamarákon keresztül szervezett workshopok, oktatóanyagok és szemináriumok segíthetnék a vállalkozások vezetőit és munkavállalóit abban, hogy megértsék, hogyan használhatják hatékonyan az AI technológiákat.
- *„Generatív AI használói jogosítvány”.* Az ECDL (Európai Számítógéphasználói Jogosítvány) mintájára egy nemzeti szintű „generatív AI-használói jogosítvány” bevezetése, amely tanúsítja az AI-alapú eszközök és platformok hatékony alkalmazását különböző üzleti és ipari környezetekben.
- *AI adoptációt elősegítő szakképzési programok.* Az AI bevezetésével kapcsolatos nem programozói szintű, hanem szakmai képesítés jellegű szerepek (pl. AI integrációs tanácsadók, AI koordinátorok) képzésének támogatása, szakmai tanfolyamok és felnőttképzési programok indítása.
- *AI adoptáció nyomon követése.* Az AI, és különösen a generatív AI elterjedésének folyamatos monitorozása, többek között az álláshirdetések figyelésével, hogy hosszú távon elérje az Excelhez hasonló penetrációt a mindennapi munkafolyamatokban. (A már említett Eurostat kísérleti mutatóhoz (Eurostat, 2023) hasonló, folyamatosan frissülő statisztikai indikátor létrehozása igen megfontolandó.)
- *AI alkalmazását támogató pályázati források.* Célzott pénzügyi ösztönzők bevezetése az AI technológiák vállalati bevezetésének elősegítésére. Pályázati források biztosítása KKV-k számára, hogy kísérleti projekteken keresztül integrálhassák az AI megoldásokat folyamataikba.
- *Állami IT erőforrások elérhetővé tétele.* Az AI kutatás és fejlesztés előmozdítása érdekében az állami szektor által biztosított IT infrastruktúrák (például számítási kapacitások, adatbázisok) hozzáférhetővé tétele vállalatok és kutatóintézetek számára.
- *Autóipari AI fejlesztések célzott támogatása.* Az autóipar az AI egyik vezető alkalmazási területe Magyarországon. Kiemelt iparpolitikai programok és célzott támogatások bevezetése az önvezető technológiák és AI-alapú mobilitási megoldások fejlesztésének ösztönzésére.

Ezek az intézkedések elősegíthetik az AI szélesebb körű alkalmazását a magyar munkaerőpiacon, növelhetik az AI-technológiák adaptációs szintjét, és biztosíthatják, hogy a generatív AI megoldások is beépüljenek a vállalatok mindennapi működésébe.

Irodalom

1. Anthropic. (2025). The Anthropic Economic Index [Accessed: 2025-02-15. Full paper available at: https://assets.anthropic.com/m/2e23255f1e84ca97/original/Economic.Tasks_AI_Paper.pdf]. <https://www.anthropic.com/news/the-anthropic-economic-index>
2. Appenzeller, G. (2024). Welcome to LLMflation – LLM inference cost is going down fast [Accessed: 2025-02-15]. <https://a16z.com/llmflation-llminference-cost/>
3. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., . . . , Liang, P. (2022). On the Opportunities and Risks of Foundation Models. <https://doi.org/10.48550/arXiv.2108.07258>
4. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., és tsai. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
5. CEPS (2025). *Solving Europe’s AI Talent Equation*. Centre for European Policy Studies. <https://cdn.ceps.eu/2025/07/solving-europes-ai-talent-equation.pdf>
6. Daas, P., Maślankowski, J., Salgado, D., Quaresma, S., Tuoto, T., Di Consiglio, L., Brancato, G., Righi, P., Six, M., & Kowarik, A. (2020. június 17). *Work package K: Methodology and quality. Deliverable K5: First draft of the methodological report (Deliverable Sz. K5) (Final version. Grant Agreement Number: 847375-2018-NL-BIGDATA). ESSnet Big Data II (Eurostat)*. https://cros.ec.europa.eu/system/files/2024-04/WPK_Deliverable_K5_First_draft_of_the_methodological_report_2020-06-17_Finalv3.pdf
7. Dechter, R. (1986). Learning While Searching in Constraint-Satisfaction-Problems. *AAAI*, 178–185.
8. Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. <https://arxiv.org/abs/2303.10130>
9. Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 226–231.
10. European Commission Joint Research Centre. (2025). *Anticipating the impact of AI on occupations: a JRC methodology*. European Commission. https://ai-watch.ec.europa.eu/news/jrc-methodology-assesses-which-occupations-face-greatest-disruption-uptake-ai-2025-06-30_en
11. Eurostat (2023, december). Labour market demand for ICT specialists in online job advertisements [Experimental statistics published in December 2023 :contentReference[oaicite:1]index=1]. <https://ec.europa.eu/eurostat/web/experimental-statistics/labour-market-demand-ict-specialists>
12. Felten, E. W., Raj, M., & Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, 108, 54–57. <https://doi.org/10.1257/pandp.20181021>
13. Felten, E. W., Raj, M., & Seamans, R. (2019). The Occupational Impact of Artificial Intelligence: Labor, Skills, and Polarization. NYU Stern School of Business. <https://doi.org/10.2139/ssrn.3368605>

14. Felten, E. W., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12), 2195–2217. <https://doi.org/10.1002/smj.3286>
15. Felten, E. W., Raj, M., & Seamans, R. (2023). Occupational Heterogeneity in Exposure to Generative AI [<https://ssrn.com/abstract=4414065>]. Available at SSRN. <https://doi.org/10.2139/ssrn.4414065>
16. Ferenczy, A. (2025). Hungary and AI: efforts and opportunities in comparison with Singapore. <https://arxiv.org/abs/2507.05280>
17. Fradkov, A. L. (2020). Early History of Machine Learning [21st IFAC World Congress]. *IFAC-PapersOnLine*, 53(2), 1385–1390. <https://doi.org/10.1016/j.ifacol.2020.12.1888>
18. Gallacher, G., & Aoki, Y. (2025). The Rise of Responsible AI Jobs. Indeed Hiring Lab. <https://www.hiringlab.org/2025/06/17/the-rise-of-responsible-ai-jobs>
19. Gmyrek, P., Berg, J., & Bescond, D. (2023). Generative AI and jobs; a global analysis of potential effects on job quantity and quality. ILO working paper; 96. <https://doi.org/10.54394/fhem8239>
20. Google (2025). Google Trends [Accessed: 2025-02-19]. <https://trends.google.com>
21. Jobspikr (2024). Jobspikr Job Market Dataset [Accessed: 2025-02-15]. <https://www.jobspikr.com/contact-us>
22. Khattab, O., Singhvi, A., Maheshwari, P., Zhang, Z., Santhanam, K., Vardhamanan, S., Haq, S., Sharma, A., Joshi, T. T., Moazam, H., Miller, H., Zaharia, M., & Potts, C. (2023). DSPy: Compiling Declarative Language Model Calls into Self-Improving Pipelines. <https://arxiv.org/abs/2310.03714>
23. Kleene, S. C. (1951, december). *Representation of Events in Nerve Nets and Finite Automata* (Research Memorandum Sz. RM-704). RAND Corporation. Santa Monica, CA. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM704.pdf
24. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2021). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. <https://arxiv.org/abs/2005.11401>
25. Li, H., Leung, J., & Shen, Z. (2024). Towards Goal-oriented Prompt Engineering for Large Language Models: A Survey. <https://arxiv.org/abs/2401.14043>
26. Lightcast (2025). Beyond the Buzz: Developing the AI Skills Employers Actually Need. <https://lightcast.io/resources/blog/beyond-the-buzzpress-release-2025-07-23>
27. Liu, H., Zhao, S., Zhang, X., Zhang, F., Wang, W., Ma, F., Chen, H., Yu, H., & Zhang, X. (2024). Liberating Seen Classes: Boosting Few-Shot and Zero-Shot Text Classification via Anchor Generation and Classification Reframing. <https://arxiv.org/abs/2405.03565>
28. McCann, B., Keskar, N. S., Xiong, C., & Socher, R. (2018). The Natural Language Decathlon: Multitask Learning as Question Answering. <https://arxiv.org/abs/1806.08730>
29. OECD (2025, június). *Emerging Divides in the Transition to Artificial Intelligence* (Policy Report). Organisation for Economic Co-operation and Development (OECD). Paris. <https://doi.org/10.1787/7376c776-en>

30. OpenAI (2025). OpenAI Platform [Accessed: 2025-02-16]. <https://platform.openai.com/>
31. Pizzinelli, C., Shibata, I., & Stavrev, E. (2023). Labor Market Exposure to AI: Cross-country Differences and Implications. *IMF Working Papers*, 2023 (216). <https://doi.org/10.5089/9798400254802.001>
32. Profession.hu Kft. (2025). Profession.hu [Accessed: 2025-07-29].
33. PwC (2025). *The Fearless Future: 2025 Global AI Jobs Barometer*. PricewaterhouseCoopers. <https://www.pwc.com/gx/en/issues/artificialintelligence/ai-jobs-barometer.html>
34. Robert Bosch Kft. (2021, március). Driver assistance systems and automated driving function development at Bosch [Press release].
35. Rogers, E. M. (2003). *Diffusion of Innovations* (5. kiad.). Free Press.
36. Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49.
37. Satopää, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011). Finding a „Kneedle” in a Haystack: Detecting Knee Points in System Behavior. arXiv preprint arXiv:1108.3085. <https://raghavan.usc.edu/papers/kneedle-simplex11.pdf>
38. Si, C., Gan, Z., Yang, Z., Wang, S., Wang, J., Boyd-Graber, J., & Wang, L. (2023). Prompting GPT-3 To Be Reliable. <https://arxiv.org/abs/2210.09150>
39. Szabados, L. (2024). AI’s effect on labour: What does economic literature say? *The International Journal of Engineering and Science*, 13(5), 328–350.
40. Szabados, L. (2025). The impact of AI tools on the work and labour market prospects of programmers. *Vezetéstudomány / Budapest Management Review*, 56(4), 42–56. <https://doi.org/10.14267/veztud.2025.04.04>
41. Wang, Z., Pang, Y., & Lin, Y. (2023). Large Language Models Are Zero-Shot Text Classifiers. <https://arxiv.org/abs/2312.01044>
42. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. <https://arxiv.org/abs/2201.11903>
43. World Economic Forum (2025). *The Future of Jobs Report 2025*. World Economic Forum. <https://www.weforum.org/publications/the-future-of-jobs-report-2025/>
44. Yeh, C.-C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., Silva, D. F., Mueen, A., & Keogh, E. J. (2016). Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 1317–1322.
45. Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2024). Can Large Language Models Transform Computational Social Science? <https://arxiv.org/abs/2305.03514>

Függelék

Reguláris kifejezés a technológiai fogalmak detektálására

Az alábbi reguláris kifejezés (RegEx) célja, hogy különböző nyelveken (angol, német, magyar) előforduló technológiai kifejezéseket azonosítsa szöveges tartalmakban. A kifejezés támogatja az egyes kifejezések eltérő helyesírási változatait, valamint bizonyos nyelvtani ragokat is figyelembe vesz magyar nyelv esetén. Az alábbi kifejezést Pythonban használtuk a technológiai kulcsszavak detektálására és a Pandas keretrendszer hatékony eszközeivel alkalmaztuk az adatsorra. A futásidő nagyságrendileg 5 perc alatt maradt.

```
pattern = r"""\b(AI|
Artificial\s+Intelligence|
Machine\s+Learning|
Computer\s+Vision|
Large\s+Language\s+Model(s)?|
Vision\s+Systems|
ChatGPT|
Natural\s+Language\s+Processing|
Data\s+Science|
Data\s+Sciences|
Künstliche\s+Intelligenz|
Maschinelles\s+Lernen|
Computersehen|
Große\s+Sprachmodelle|
Natürliche\s+Sprachverarbeitung|
Sprachverarbeitung|
Datenwissenschaft|
Datenwissenschaften|

mesterséges\s+intelligencia|
gépi\s+tanulás|
számítógépes\s+látás|
gépi\s+látás|
vision\s+rendszer(ek)?|
nagy\s+nyelvi\s+modell(ek)?|
természetes\s+nyelv(i?)\s+feldolgozás|
nyelv(i?)\s+feldolgozás|
nyelvfeldolgozás|
adattudomány(ok)?|
adat\s+tudomány(ok)?|

Aritificial\s+Inteligenace|
Machin\s+Lerning|
NLP|
LLMs?
)\b(?:val|vel|ban|ben|nak|nek|
ból|ből|hoz|hez|höz|
tól|től|ra|re|al|el|s|ok)?"""
```

Ez a reguláris kifejezés az alábbi tulajdonságokkal rendelkezik:

- Több nyelven képes felismerni a releváns kifejezéseket (angol, német, magyar).
- Figyelembe veszi az egyes nyelvekre jellemző helyesírási variációkat, rövidítéseket és elgépeléseket (például: AI, NLP, LLMs).
- A magyar nyelv esetén felismeri a technológiai kifejezésekhez kapcsolódó ragozott formákat is.
- Pythonban történő használatra optimalizált, a re.IGNORECASE és a re.VERBOSE flagekkel együtt alkalmazható.

A reguláris kifejezést sikeresen alkalmaztuk különböző szövegeken technológiai kifejezések detektálására, lehetővé téve azok szisztematikus azonosítását és elemzését.

DSPy optmilaizált prompt szemantikus osztályozásra

Az optimalizációs eljárás eredményeként létrejött végső, a tesztadaton hibátlan teljesítményű prompt a következő volt:

```
Conduct a thorough examination of the provided job
description to determine its association with AI and ML.
Classify it as 'True' if any explicit mentions are noted
regarding fields such as Artificial Intelligence,
Machine Learning, Data Science, Deep Learning, or
related concepts such as Neural Networks, Natural
Language Processing.
Also consider subtle hints or phrasing that suggest
involvement with such technologies, and be mindful
of variations in language from across the globe.
If these elements are absent, unequivocally mark it
as 'False'. Please respond with only the two words
'True' or 'False', maintaining accuracy and clarity
while avoiding additional commentary. }
```

„AI IN THE MARKET”: ADOPTION OF ARTIFICIAL INTELLIGENCE
SYSTEMS IN LIGHT OF LARGE-SCALE ANALYSIS OF HUNGARIAN
JOB ADVERTISEMENTS

The widespread application of Artificial Intelligence (AI) systems in business is reshaping the job market, yet the extent to which this technological potential has translated into practical adoption remains an open question. Our study examines the presence of AI-related job opportunities in Hungary by analyzing over 1.3 million online job advertisements collected between 2019 and 2024. Our goal is to provide a data-driven assessment of the AI labor market in Hungary, analyze its temporal dynamics, sectoral adoption, and identify emerging corporate strategies.

AI systems have undergone multiple waves of development, from the early symbolic expert systems of the 1950s to the dominance of Machine Learning (ML) in the early 2000s. A major breakthrough occurred in 2012 with Deep Learning proving its superiority in various applications. The next paradigm shift came after 2020 with the emergence of Large Language Models (LLMs), which demonstrated broad problem-solving capabilities and led to an unprecedented surge in AI-driven automation. The global excitement surrounding OpenAI's ChatGPT in 2022 significantly increased interest in AI applications across industries.

To assess AI penetration in the Hungarian labor market, we conducted a computational analysis of a large-scale dataset of job postings. The dataset, sourced from the JobsPikr platform, includes advertisements in Hungarian, English, and German. We applied Natural Language Processing (NLP) techniques to categorize job postings according to AI relevance. Two classification methods were employed: (1) a keyword-based approach using Regular Expressions (RegEx) to identify explicit mentions of AI-related terms, and (2) a more advanced semantic classification using Large Language Models (LLM-based DSPy optimization) to assess the actual AI-related nature of the jobs.

Our findings reveal that AI-related job postings account for only 1.34% of all job advertisements in Hungary, indicating a relatively low market penetration of AI-driven job roles. The distribution of AI-related postings is heavily concentrated in Budapest, aligning with the city's status as the country's primary technology hub. Two notable waves of AI-related hiring activity were observed: one in 2020–2021, primarily driven by the automotive industry (self-driving vehicle research and development), and a second, smaller increase in 2023, likely influenced by the growing awareness of generative AI solutions.

Sectoral analysis indicates that the majority of AI-related job postings originate from the IT sector (24%), followed by consulting (11%) and financial services. Automotive companies such as Bosch and Continental, along with AI-focused firms like AIMotive, play a dominant role in AI-related hiring, particularly in the domain of autonomous vehicle development. Different corporate strategies for AI adoption were identified: (1) short-term „AI hiring waves” focused on building fixed AI capabilities, (2) gradual AI integration with a sustained hiring trend, and (3) AI branding strategies where AI is mentioned in all job advertisements to enhance corporate image, even when the job itself is unrelated to AI.

The latest AI technologies, such as generative AI and ChatGPT, have only marginally entered the Hungarian job market. With just 128 explicit mentions in job postings, the adoption of these tools remains limited. In contrast, more traditional AI roles such as software engineers, machine learning specialists, and data scientists dominate the AI job landscape. Methodologically, our study demonstrates that the LLM-based DSPy classification significantly outperforms the keyword-based RegEx approach by reducing false positives and accurately identifying AI-related roles. This highlights the necessity of semantic analysis when assessing the adoption of AI in job markets.

The study presents key policy recommendations for accelerating AI adoption in Hungary, including (1) AI literacy programs, (2) targeted AI workforce training, (3) financial incentives for AI technology adoption, (4) public-sector computing infrastructure support, and (5) sector-specific initiatives, particularly in the automotive industry, which has proven to be a leading driver of AI job growth. As AI technologies become more integrated into business operations, explicit mentions of AI in job postings may decline, mirroring the adoption patterns of earlier general-purpose technologies like office software. Future research should explore alternative approaches for measuring AI adoption in workplaces, such as occupational task analysis and employer surveys. Overall, our findings suggest that while AI adoption is steadily increasing in Hungary, it remains concentrated in specific industries and urban centers, with a slow diffusion into the broader economy. The extent to which generative AI solutions will reshape the Hungarian job market in the coming years remains an open question.