MKD revue, ročník 12 (2025), číslo 2, ISSN: 1339-777X

Od symbolického uvažovania po generatívne modely: Historický a konceptuálny vývoj umelej inteligencie

Vladimír FILIP

Abstrakt

Článok sa zameriava na historický a konceptuálny vývoj umelej inteligencie so zvláštnym dôrazom na prechod od symbolických a pravidlami riadených prístupov k sub-symbolickým, dátovo orientovaným a generatívnym modelom. Vychádza z predpokladu, že súčasné formy umelej inteligencie predstavujú vyústenie dlhodobej intelektuálnej tradície, v ktorej sa postupne formovali predstavy o efektívnych metódach, formalizácii myslenia a mechanizácii kognitívnych procesov. Text sleduje kľúčové myšlienkové línie od raných filozofických a matematických úvah o algoritmickosti, cez formálne modely výpočtu Alonza Churcha a Alana Turinga, vznik umelej inteligencie ako vedeckej disciplíny a rozvoj symbolickej paradigmy, až po jej kritiku a nástup konektivistických a generatívnych prístupov. Metodologicky je článok koncipovaný ako teoreticko-konceptuálna štúdia založená na historicko-genetickej analýze a komparácii vybraných výskumných paradigiem umelej inteligencie. Cieľom článku je prispieť k hlbšiemu porozumeniu súčasných generatívnych modelov v kontexte ich historického, epistemologického a filozofického vývoja.

Kľúčové slová

umelá inteligencia, symbolická umelá inteligencia, výpočtové formalizmy, strojové učenie, generatívne modely

MKD revue, ročník 12 (2025), číslo 2, ISSN: 1339-777X

Abstract

The article focuses on the historical and conceptual development of artificial intelligence, with particular emphasis on the transition from symbolic and rule-based approaches to sub-symbolic, data-oriented, and generative models. It is based on the assumption that current forms of artificial intelligence represent the culmination of a long intellectual tradition in which ideas about effective methods, the formalization of thinking, and the mechanization of cognitive processes have gradually taken shape. The text follows key lines of thought from early philosophical and mathematical reflections on algorithmicity, through the formal models of computation by Alonzo Church and Alan Turing, the emergence of artificial intelligence as a scientific discipline and the development of the symbolic paradigm, to its criticism and the advent of connectivist and generative approaches. Methodologically, the article is conceived as a theoretical-conceptual study based on historical-genetic analysis and comparison of selected research paradigms of artificial intelligence. The aim of the article is to contribute to a deeper understanding of current generative models in the context of their historical, epistemological, and philosophical development.

Keywords

artificial intelligence, symbolic artificial intelligence, computational formalisms, machine learning, generative models

Úvod

Umelej inteligencii sa v súčasnosti venuje mimoriadna pozornosť nielen v oblasti informatiky a technických vied, ale aj v rámci filozofie, kognitívnej vedy, mediálnych a kultúrnych štúdií. Rýchly rozvoj dátovo orientovaných prístupov, hlbokého učenia a generatívnych modelov viedol k zásadným zmenám v spôsobe, akým sú inteligentné systémy navrhované, implementované a interpretované. Tento vývoj však zároveň vyvoláva otázky týkajúce sa samotnej povahy inteligencie, vzťahu medzi výpočtom a porozumením, ako aj kontinuity medzi

historickými koncepciami myslenia a ich súčasnými technickými realizáciami. Hoci sa súčasná diskusia o umelej inteligencii často sústreďuje na praktické aplikácie a etické dôsledky jej využívania, menej pozornosti sa venuje historickému a konceptuálnemu vývoju základných ideí, ktoré umožnili jej vznik. Moderné systémy umelej inteligencie totiž nevznikli vo vákuu, ale predstavujú vyústenie dlhodobej intelektuálnej tradície, v ktorej sa postupne formovali predstavy o efektívnych metódach, formalizácii myslenia a mechanizácii kognitívnych procesov.

Cieľom tohto článku je analyzovať historický a konceptuálny vývoj umelej inteligencie so zameraním na prechod od symbolických a pravidlami riadených prístupov k sub-symbolickým, dátovo orientovaným a generatívnym modelom. Text sleduje vývoj kľúčových myšlienkových línií, ktoré viedli od raných predstáv o efektívnych metódach a vypočítateľnosti, formulovaných v prácach Churcha a Turinga, až po súčasné modely založené na hlbokom učení a štatistickej analýze veľkých dátových súborov. Dôraz sa pritom kladie na pochopenie toho, ako sa menili predstavy o reprezentácii poznania, učení a význame v rámci rôznych paradigmatických prístupov umelej inteligencie.

Metodológia výskumu

Predkladaný článok je koncipovaný ako teoreticko-konceptuálna štúdia zameraná na historickú a analytickú reflexiu vývoja umelej inteligencie. Metodologický rámec práce vychádza z kombinácie historicko-genetického prístupu a konceptuálnej analýzy, ktorých cieľom je identifikovať a interpretovať kľúčové myšlienkové línie, paradigmatické zlomy a teoretické predpoklady, ktoré formovali vývoj umelej inteligencie od raných úvah o efektívnych metódach až po súčasné generatívne modely. Základnou metódou výskumu je kritická analýza primárnych a sekundárnych zdrojov. Medzi primárne zdroje patria pôvodné práce autorov, ktorí zásadným spôsobom ovplyvnili formovanie teórie výpočtu a umelej inteligencie, najmä texty Alonza Churcha, Alana Turinga, Johna McCarthyho, Allena Newella a Herberta A. Simona. Sekundárne zdroje tvoria prehľadové monografie a syntetické štúdie z oblasti umelej

inteligencie, filozofie informácie a kognitívnej vedy, ktoré umožňujú kontextualizovať historické texty v širšom interdisciplinárnom rámci.

Analytický postup článku je založený na komparatívnom porovnávaní vybraných výskumných paradigiem umelej inteligencie, najmä symbolického a sub-symbolického prístupu. Pozornosť sa pritom sústreďuje na spôsob reprezentácie poznania, mechanizmy učenia a vysvetľovania inteligentného správania v rámci jednotlivých modelov. Cieľom tejto komparácie nie je hodnotenie technickej výkonnosti konkrétnych algoritmov, ale identifikácia ich epistemologických predpokladov a teoretických dôsledkov pre chápanie inteligencie ako takej.

Výskum má prevažne kvalitatívny charakter a nepracuje s empirickými dátami ani experimentálnymi výsledkami. Text sa nesnaží o vyčerpávajúci prehľad všetkých existujúcich prístupov k umelej inteligencii, ale cielene sa zameriava na reprezentatívne modely a koncepcie, ktoré mali zásadný vplyv na ďalší vývoj odboru. Výber zdrojov je vedený kritériami vedeckej relevantnosti, citačného ohlasu a dlhodobej významnosti v rámci odbornej diskusie. Metodologické obmedzenia článku spočívajú najmä v jeho teoretickom zameraní, ktoré neumožňuje priame overovanie záverov prostredníctvom empirického výskumu. Napriek tomu však zvolený metodologický prístup umožňuje hlbšiu reflexiu konceptuálnych základov umelej inteligencie a prispieva k porozumeniu súčasných generatívnych modelov v kontexte ich historického a filozofického vývoja. Metodológia tak vytvára rámec pre interpretáciu umelej inteligencie nielen ako technického nástroja, ale aj ako kultúrneho a epistemického fenoménu.

Efektívne metódy a počiatky formalizovaného uvažovania

Myšlienka, že ľudské uvažovanie možno opísať prostredníctvom presne definovaných postupov, má hlboké historické korene, ktoré predchádzajú vzniku modernej informatiky. Už v staroveku sa v oblasti aritmetiky a geometrie objavovali systematické postupy riešenia problémov, ktoré umožňovali opakovateľné dosahovanie výsledkov. Tieto postupy možno spätne chápať ako rané formy efektívnych metód, teda konečných a jednoznačných pravidiel vedúcich k riešeniu určitej triedy úloh (Knuth, 1997).

Systematická reflexia povahy takýchto metód sa začala rozvíjať najmä v kontexte filozofických úvah o poznaní a racionalite. V stredoveku sa objavujú pokusy o formalizáciu myslenia, ktorých cieľom bolo zredukovať zložité otázky na kombinácie elementárnych pojmov. Významným príkladom je kombinatorický systém Ramona Llulla, ktorý sa usiloval vytvoriť univerzálny mechanizmus generovania a hodnotenia výrokov prostredníctvom konečného súboru symbolov a pravidiel ich kombinovania (Yates, 1999). Hoci bol tento systém primárne motivovaný teologicky, obsahoval dôležitú epistemologickú intuície, že proces uvažovania možno formalizovať nezávisle od konkrétneho obsahu myslenia.

V novovekej filozofii sa tento prístup ďalej rozvíjal v súvislosti s mechanickým chápaním sveta. Myslenie začalo byť interpretované ako proces manipulácie s reprezentáciami podľa presných pravidiel. Thomas Hobbes v tomto kontexte výslovne postuloval, že rozumové operácie možno chápať ako formu výpočtu, pričom uvažovanie prirovnával k aritmetickému sčítaniu a odčítaniu pojmov (Hobbes, 1996). Tento pohľad predstavoval zásadný posun, pretože mentálne procesy prestali byť chápané výlučne ako introspektívne alebo kvalitatívne javy a začali byť analyzované ako potenciálne formalizovateľné operácie.

Kľúčovým spoločným prvkom týchto historických koncepcií je implicitná predstava algoritmickosti, teda presvedčenie, že existujú konečné, determinované a vykonateľné postupy, ktorých aplikácia vedie k riešeniu problémov. Hoci pojem algoritmu v dnešnom technickom zmysle ešte nebol explicitne definovaný, základná intuícia o existencii efektívnych výpočtových postupov bola prítomná už v týchto raných úvahách. Práve táto intuícia sa v 20. storočí stala východiskom pre formálne uchopenie pojmu výpočtu a vypočítateľnosti (Copeland, 2004).

Efektívne metódy tak možno chápať ako historický predobraz moderných výpočtových modelov. Ich význam nespočíva iba v praktickej použiteľnosti, ale najmä v tom, že postupne formovali predstavu o myslení ako o procese, ktorý je možné analyzovať a modelovať nezávisle od jeho biologického nositeľa. Táto predstava vytvorila konceptuálne východisko pre matematickú teóriu výpočtu a neskorší vznik umelej inteligencie ako samostatnej vedeckej disciplíny (Floridi, 2011).

Výpočtové formalizmy: Turing, Church a hranice vypočítateľnosti

Zásadný obrat v chápaní efektívnych metód nastal v prvej polovici 20. storočia, keď sa otázka vypočítateľnosti stala predmetom presnej matematickej analýzy. Dovtedajšie intuitívne predstavy o algoritmickom postupe boli nahradené snahou o formálnu definíciu pojmu výpočtu, ktorá by bola nezávislá od konkrétnej implementácie či fyzického realizátora. V tomto kontexte zohrali kľúčovú úlohu práce Alonza Churcha a Alana Turinga, ktorí nezávisle od seba navrhli formálne modely výpočtu schopné zachytiť intuitívny pojem efektívne vykonateľného postupu (Church, 1936; Turing, 1936). Churchov prístup vychádzal z lambda kalkulu, formálneho systému založeného na abstrakcii a aplikácii funkcií. Lambda kalkulus umožnil presne definovať triedu funkcií, ktoré možno považovať za efektívne vypočítateľné, pričom manipulácia so symbolmi bola redukovaná na niekoľko jednoduchých transformačných pravidiel (Kleene, 1952). Napriek svojej zdanlivej jednoduchosti sa lambda kalkulus ukázal ako mimoriadne expresívny formálny systém, schopný reprezentovať širokú škálu výpočtových procesov.

Nezávisle od Churcha navrhol Alan Turing vlastný model výpočtu v podobe abstraktného stroja, dnes známeho ako Turingov stroj. Tento model pozostáva z nekonečnej pásky rozdelenej na bunky, hlavy schopnej čítať a zapisovať symboly a konečného riadiaceho automatu určujúceho správanie stroja v jednotlivých krokoch výpočtu. Turingov prínos spočíval v tom, že intuitívny pojem mechanického výpočtu previedol do presne definovanej formálnej štruktúry, ktorá umožnila matematickú analýzu samotného procesu výpočtu (Turing, 1936).

Ako sa neskôr ukázalo, formálne systémy navrhnuté Churchom a Turingom sú výpočtovo ekvivalentné. Každá funkcia, ktorú možno definovať v lambda kalkule, je realizovateľná pomocou Turingovho stroja, a naopak. Táto ekvivalencia viedla k formulácii tzv. Churchovej–Turingovej tézy, podľa ktorej každá intuitívne efektívne vypočítateľná funkcia je vypočítateľná Turingovým strojom (Copeland, 2004). Hoci nejde o matematickú vetu v striktnom zmysle slova, ale o filozofickú tézu, jej význam pre teóriu výpočtu a informatiku je zásadný.

Formálne modely výpočtu však neumožnili len presnejšie definovať pojem algoritmu, ale aj identifikovať hranice toho, čo je možné vypočítať. Turing vo svojej práci demonštroval

existenciu neriešiteľných problémov, medzi ktoré patrí napríklad problém zastavenia. Ukázal, že neexistuje všeobecný algoritmus, ktorý by pre ľubovoľný program a vstup dokázal rozhodnúť, či sa daný program zastaví alebo bude pokračovať v nekonečnom výpočte (Turing, 1936). Tento výsledok mal hlboké epistemologické dôsledky, pretože naznačil, že existujú presne formulované otázky, na ktoré nie je možné algoritmicky získať odpoveď. Význam lambda kalkulu a Turingovho stroja presahuje rámec čistej matematiky. Lambda kalkulus sa stal teoretickým základom viacerých programovacích jazykov, najmä jazyka Lisp, ktorý zohral kľúčovú úlohu v ranom vývoji umelej inteligencie. Práca s funkciami ako s dátami, dynamická tvorba výrazov a schopnosť manipulovať vlastným kódom umožnili experimentovanie s modelmi myslenia a učenia v symbolickej AI (McCarthy, 1960).

Turingov model výpočtu zároveň položil základy modernej informatiky ako disciplíny, ktorá sa nezaoberá len konkrétnymi strojmi, ale všeobecnými princípmi spracovania informácií. Presun otázok myslenia a inteligencie z filozofického rámca do výpočtového kontextu vytvoril predpoklady pre vznik umelej inteligencie ako samostatnej oblasti výskumu. Výpočtové formalizmy Churcha a Turinga tak predstavujú konceptuálny most medzi historickými úvahami o mechanizácii myslenia a modernými snahami o jeho simuláciu pomocou algoritmických systémov.

Zrodenie umelej inteligencie ako vedeckej disciplíny

Hoci výpočtové formalizmy Alonza Churcha a Alana Turinga poskytli presnú definíciu pojmu výpočtu a jeho hraníc, samy osebe ešte neimplikovali vznik umelej inteligencie ako samostatnej oblasti výskumu. Rozhodujúcim krokom bol presun pozornosti od abstraktnej otázky vypočítateľnosti k ambicióznej snahe simulovať aspekty ľudskej inteligencie pomocou výpočtových systémov. Tento posun sa explicitne prejavil v polovici 20. storočia, keď sa začala formovať predstava, že mentálne procesy možno nielen analyzovať, ale aj implementovať na strojoch.

Za symbolický začiatok umelej inteligencie ako vedeckej disciplíny sa tradične považuje letný výskumný projekt uskutočnený v roku 1956 na Dartmouth College. Iniciátorom tohto podujatia

bol John McCarthy, ktorý vo svojom návrhu definoval cieľ projektu ako snahu vychádzať z predpokladu, že „každý aspekt učenia alebo akákoľvek iná vlastnosť inteligencie môže byť opísaná tak presne, že je možné zostrojiť stroj, ktorý ju dokáže simulovať“ (McCarthy et al., 1955). Táto formulácia predstavovala výrazný epistemologický záväzok, keďže implicitne predpokladala, že inteligencia je v princípe formalizovateľná a algoritmizovateľná. Účastníci dartmouthského workshopu, medzi ktorými figurovali mená ako Marvin Minsky, Claude Shannon či Herbert A. Simon, reprezentovali rôzne vedecké disciplíny, od matematiky a logiky až po psychológiu a elektrotechniku. Spoločným menovateľom ich prístupov bola viera v to, že digitálny počítač, chápaný ako univerzálny výpočtový stroj v Turingovom zmysle, poskytuje vhodnú platformu na experimentovanie s modelmi inteligentného správania (Russell a Norvig, 2021). V tomto kontexte sa umelá inteligencia začala formovať ako interdisciplinárny výskumný program, ktorý prepájal formálne teórie s praktickými implementáciami.

Raný výskum umelej inteligencie bol výrazne ovplyvnený symbolickým prístupom. Inteligencia bola chápaná predovšetkým ako schopnosť manipulovať symbolmi podľa explicitne definovaných pravidiel, pričom dôraz sa kládol na riešenie problémov, dokazovanie viet a spracovanie jazyka. Tento prístup bol teoreticky ukotvený v tzv. fyzikálno-symbolickej hypotéze mysle, podľa ktorej je každý systém schopný všeobecnej inteligentnej činnosti nevyhnutne systémom fyzickej manipulácie symbolov (Newell a Simon, 1976). Výpočtové modely sa tak stali nielen nástrojmi, ale aj explanatórnymi modelmi kognitívnych procesov.

Zároveň sa však už v tomto období začali objavovať napätia medzi ambicióznymi cieľmi a reálnymi možnosťami vtedajších systémov. Prvé úspechy v oblasti riešenia formálnych úloh viedli k optimistickým predpovediam o rýchlom dosiahnutí všeobecnej umelej inteligencie, ktoré sa však ukázali ako predčasné. Obmedzený výpočtový výkon, nedostatočné znalostné bázy a problémy so škálovateľnosťou symbolických systémov postupne poukázali na limity tohto prístupu. Napriek tomu dartmouthský projekt položil trvalé základy umelej inteligencie ako samostatnej disciplíny, keď jasne vymedzil jej predmet skúmania, metodologické ambície a vzťah k výpočtovej teórii.

Vznik umelej inteligencie tak možno chápať ako historický moment, v ktorom sa teoretické úvahy o vypočítateľnosti spojili s praktickými snahami o modelovanie inteligentného správania. Tento moment vytvoril rámec, v ktorom sa následne rozvinuli rôzne výskumné paradigmy, od symbolickej AI až po súčasné dátovo orientované a generatívne modely.

Symbolická umelá inteligencia a paradigma GOFAI

Po inštitucionalizácii umelej inteligencie ako samostatnej oblasti výskumu sa v nasledujúcich desaťročiach etabloval prístup, ktorý je dnes retrospektívne označovaný ako symbolická umelá inteligencia alebo Good Old-Fashioned Artificial Intelligence (GOFAI). Tento prístup vychádzal z predpokladu, že inteligencia je v podstate proces manipulácie symbolov na základe explicitne definovaných pravidiel. Symboly v tomto kontexte reprezentovali objekty, vlastnosti alebo vzťahy vo svete a ich spracovanie malo prebiehať prostredníctvom logických a formálnych operácií.

Teoretickým jadrom symbolickej AI bola fyzikálno-symbolická hypotéza mysle, ktorú formulovali Allen Newell a Herbert A. Simon. Podľa tejto hypotézy je každý systém schopný všeobecnej inteligentnej činnosti nevyhnutne systémom fyzickej manipulácie symbolov a naopak, každý dostatočne komplexný systém manipulujúci symbolmi má potenciál vykazovať inteligentné správanie (Newell a Simon, 1976). Tento názor poskytol silné epistemologické opodstatnenie pre vývoj programov, ktoré mali modelovať ľudské riešenie problémov, uvažovanie a rozhodovanie. Symbolická AI kládla dôraz najmä na oblasť riešenia problémov, plánovania a dokazovania viet. Typickým príkladom sú programy ako Logic Theorist alebo General Problem Solver, ktoré sa snažili formalizovať ľudské heuristiky riešenia úloh pomocou vyhľadávania v priestore stavov (Newell a Simon, 1956). Inteligencia bola v tomto rámci chápaná ako schopnosť systematicky prehľadávať priestor možných riešení a aplikovať vhodné pravidlá na dosiahnutie cieľa.

Kľúčovým aspektom symbolického prístupu bola reprezentácia poznania. Znalosti o svete boli ukladané v explicitnej forme, často prostredníctvom logických výrokov, produkčných pravidiel alebo sémantických sietí. Tento spôsob reprezentácie umožňoval transparentné a

interpretovateľné uvažovanie, keďže každý krok inferencie bol v princípe sledovateľný a vysvetliteľný. Práve táto vlastnosť bola považovaná za jednu z hlavných výhod symbolickej AI v porovnaní s neskoršími sub-symbolickými prístupmi (Russell a Norvig, 2021).

Významnou aplikáciou symbolickej AI sa stali expertné systémy, ktoré sa snažili zachytiť znalosti ľudských expertov v úzko vymedzených oblastiach, ako bola medicína alebo technická diagnostika. Tieto systémy využívali rozsiahle bázy pravidiel a inferenčné mechanizmy na poskytovanie odporúčaní alebo rozhodnutí. V 70. a 80. rokoch boli expertné systémy považované za jednu z najúspešnejších praktických realizácií umelej inteligencie a významne prispeli k jej rozšíreniu v priemyselnej praxi (Buchanan a Shortliffe, 1984). Napriek počiatočným úspechom sa však postupne ukázali aj zásadné limity symbolického prístupu. Symbolické systémy sa ukázali ako krehké, ťažko škálovateľné a silne závislé od kvality a úplnosti explicitne zadanej znalostnej bázy. Problémy ako kombinatorická explózia, nedostatočná schopnosť učiť sa z dát a ťažkosti s reprezentáciou neúplného alebo neurčitého poznania poukázali na obmedzenia modelu, ktorý staval inteligenciu výlučne na manipulácii symbolov (Dreyfus, 1972). Tieto nedostatky postupne viedli k hľadaniu alternatívnych prístupov, ktoré by dokázali lepšie zachytiť adaptívny a kontextuálny charakter inteligentného správania.

Symbolická umelá inteligencia tak predstavuje dôležitú, no historicky ohraničenú etapu vývoja AI. Jej prínos spočíva nielen v konkrétnych technických riešeniach, ale aj v tom, že explicitne formulovala otázku vzťahu medzi reprezentáciou, výpočtom a inteligenciou. Práve tieto otázky sa stali východiskom pre neskoršiu kritiku symbolizmu a pre vznik sub-symbolických a dátovo orientovaných prístupov.

Kritika symbolického prístupu a obrat k sub-symbolickým modelom

Napriek počiatočnému optimizmu a čiastkovým úspechom symbolickej umelej inteligencie sa už v 60. a 70. rokoch začali objavovať systematické kritiky, ktoré poukazovali na jej teoretické aj praktické obmedzenia. Tieto kritiky vychádzali z rôznych disciplín, predovšetkým z filozofie

mysle, kognitívnej vedy a epistemológie, a spochybňovali základný predpoklad, že inteligencia môže byť adekvátne vysvetlená ako manipulácia symbolov na základe formálnych pravidiel.

Jednou z najvplyvnejších filozofických kritík symbolickej AI bola argumentácia Huberta L. Dreyfusa, ktorý poukázal na to, že ľudské myslenie a konanie sú hlboko zakotvené v telesnej a situačnej skúsenosti. Podľa Dreyfusa symbolické systémy nedokážu zachytiť implicitné, neformálne a kontextovo viazané aspekty ľudskej inteligencie, ktoré nie sú ľahko formalizovateľné v podobe explicitných pravidiel (Dreyfus, 1972). Tento problém sa prejavoval najmä v situáciách, kde bolo potrebné pracovať s neúplnými informáciami, nejednoznačnými významami alebo rýchlo sa meniacim kontextom. Ďalším zásadným filozofickým zásahom do diskusie o povahe umelej inteligencie bol tzv. argument čínskej izby, ktorý formuloval John Searle. Searle týmto myšlienkovým experimentom spochybnil tvrdenie, že syntaktická manipulácia symbolov je postačujúca na vznik porozumenia a významu. Aj keď systém môže vykazovať správanie, ktoré je zvonka nerozoznateľné od inteligentného, neznamená to podľa Searla, že skutočne „rozumie“ spracovávaným symbolom (Searle, 1980). Tento argument poukázal na zásadný rozdiel medzi formálnou správnosťou výpočtu a sémantickým porozumením.

Okrem filozofických námietok sa čoraz výraznejšie prejavovali aj technické a praktické limity symbolických systémov. Jedným z najznámejších problémov bol tzv. rámcový problém, ktorý poukazuje na ťažkosti spojené s reprezentáciou a aktualizáciou poznania v dynamickom prostredí. Symbolické systémy mali problém rozhodnúť, ktoré informácie sú v danom kontexte relevantné a ktoré možno ignorovať, čo viedlo k explozívnemu nárastu počtu pravidiel a inferenčných krokov (McCarthy a Hayes, 1969). Tento jav výrazne obmedzoval škálovateľnosť symbolických prístupov.

S rastúcou komplexitou úloh sa tiež ukázalo, že ručné vytváranie a udržiavanie rozsiahlych znalostných báz je nielen časovo náročné, ale aj epistemologicky problematické. Znalosti expertov sa často ukazovali ako neúplné, nekonzistentné alebo ťažko formalizovateľné. Symbolická AI tak narážala na hranice nielen výpočtovej kapacity, ale aj samotného procesu explicitnej reprezentácie poznania (Winograd a Flores, 1986).

Tieto kritiky neviedli k okamžitému opusteniu symbolického prístupu, no postupne vytvorili intelektuálne prostredie, v ktorom sa začali presadzovať alternatívne modely inteligencie. Pozornosť sa presunula od explicitnej manipulácie symbolov k sub-symbolickým procesom, ktoré kládli dôraz na učenie z dát, adaptáciu a emergentné správanie. Inteligencia sa v tomto novom rámci prestala chápať ako striktne pravidlami riadený proces a začala byť interpretovaná ako dynamický jav vznikajúci z interakcie jednoduchších komponentov. Kritika symbolickej umelej inteligencie tak nepredstavovala jej úplné zavrhnutie, ale skôr historický moment reflexie, ktorý otvoril cestu k novým paradigmatickým prístupom. Práve v tomto kontexte sa začali formovať konektivistické modely a neskôr aj metódy strojového učenia, ktoré sa usilovali prekonať obmedzenia symbolických systémov bez toho, aby sa vzdali výpočtového rámca ako takého.

Konektivizmus, strojové učenie a generatívne modely ako nová paradigma umelej inteligencie

Postupná kritika symbolickej umelej inteligencie viedla v druhej polovici 20. storočia k hľadaniu alternatívnych modelov, ktoré by dokázali lepšie zachytiť adaptívny, kontextuálny a skúsenostný charakter inteligentného správania. Jedným z najvýznamnejších smerov, ktoré sa v tomto období začali presadzovať, bol konektivizmus, teda prístup inšpirovaný štruktúrou a fungovaním biologických neurónových sietí. Na rozdiel od symbolickej AI, ktorá pracovala s explicitnými reprezentáciami a pravidlami, konektivistické modely kládli dôraz na distribuované spracovanie informácií a učenie prostredníctvom úpravy váh medzi jednoduchými výpočtovými jednotkami (Rumelhart a McClelland, 1986).

Historické korene konektivizmu možno vystopovať už k modelu formálneho neurónu, ktorý navrhli McCulloch a Pitts v 40. rokoch 20. storočia. Tento model ukázal, že aj jednoduché výpočtové jednotky, spojené do siete, môžu realizovať logické funkcie a vykazovať komplexné správanie (McCulloch a Pitts, 1943). Napriek tomuto sľubnému začiatku však výskum neurónových sietí v 60. rokoch narazil na vážne obmedzenia, ktoré vyústili do dočasného útlmu konektivistických prístupov. Kritika perceptrónov, najmä poukázanie na ich neschopnosť riešiť

nelineárne separovateľné problémy, viedla k presvedčeniu, že neurónové siete nepredstavujú životaschopnú alternatívu k symbolickej AI (Minsky a Papert, 1969).

Obrat nastal v 80. rokoch s opätovným objavením algoritmu spätného šírenia chyby, ktorý umožnil efektívne trénovanie viacvrstvových neurónových sietí. Tento technický pokrok obnovil záujem o konektivistické modely a otvoril cestu k chápaniu učenia ako procesu postupnej adaptácie váh na základe skúsenosti s dátami (Rumelhart, Hinton a Williams, 1986). Inteligencia v tomto rámci prestala byť chápaná ako aplikácia vopred daných pravidiel a začala byť interpretovaná ako emergentný jav vznikajúci z interakcie jednoduchých výpočtových prvkov.

Rozvoj konektivizmu bol úzko spätý so širším posunom smerom k dátovo orientovaným prístupom, ktoré sa dnes súhrnne označujú ako strojové učenie. Namiesto explicitného programovania správania systému sa pozornosť presunula na návrh algoritmov, ktoré dokážu generalizovať zo skúsenosti a adaptovať sa na nové situácie. Tento posun predstavoval zásadnú zmenu paradigmy, keďže znalosti už neboli reprezentované v symbolickej podobe, ale implicitne zakódované v parametroch modelu (Mitchell, 1997).

Ďalší kvalitatívny skok nastal v prvej dekáde 21. storočia s nástupom hlbokého učenia, teda neurónových sietí s veľkým počtom vrstiev, schopných učiť sa hierarchické reprezentácie dát. Kombinácia rastúceho výpočtového výkonu, dostupnosti rozsiahlych dátových súborov a architektonických inovácií viedla k dramatickému zlepšeniu výkonu v oblastiach, ako je rozpoznávanie obrazu, spracovanie reči či analýza prirodzeného jazyka (LeCun, Bengio a Hinton, 2015). Hlboké učenie tak postupne nahradilo symbolické prístupy v mnohých praktických aplikáciách umelej inteligencie.

V posledných rokoch sa do popredia dostali generatívne modely, ktoré predstavujú ďalší krok v evolúcii dátovo orientovanej AI. Na rozdiel od diskriminačných modelov, ktorých cieľom je priraďovať vstupy k výstupným triedam, generatívne modely sa usilujú zachytiť štatistickú štruktúru dát ako takých a generovať nové, dovtedy nevidené príklady. Medzi najvýznamnejšie generatívne architektúry patria variačné autoenkódery, generatívne adversariálne siete a v

poslednom období najmä modely založené na architektúre transformera (Goodfellow et al., 2014; Vaswani et al., 2017).

Veľké jazykové modely, ktoré využívajú transformerovú architektúru, predstavujú kulmináciu tohto vývoja v oblasti spracovania prirodzeného jazyka. Jazyk je v nich modelovaný ako pravdepodobnostná štruktúra, v ktorej sa význam neobjavuje v podobe explicitných symbolických reprezentácií, ale emerguje zo vzťahov medzi tokenmi v rozsiahlych dátových korpusoch. Tento prístup zásadne spochybňuje tradičné predstavy o reprezentácii významu a opätovne otvára filozofické otázky týkajúce sa porozumenia, intencionality a povahy inteligencie (Bender et al., 2021).

Z historického hľadiska je však dôležité zdôrazniť, že generatívne modely nepredstavujú popretie výpočtovej teórie mysle ani Churchovej–Turingovej tézy. Naopak, fungujú plne v rámci výpočtového paradigmatu a realizujú výpočty, ktoré sú v princípe Turingovo vypočítateľné. Ich odlišnosť spočíva v spôsobe, akým sú znalosti reprezentované a získavané, nie v prekročení formálnych hraníc výpočtu. Generatívne modely tak možno chápať ako historicky konzistentné pokračovanie vývoja umelej inteligencie, v ktorom sa pôvodná idea mechanizácie myslenia transformovala do podoby štatistického učenia z dát.

Spojením konektivistických princípov, hlbokého učenia a generatívnych architektúr vzniká nový konceptuálny rámec umelej inteligencie, ktorý kladie dôraz na adaptáciu, generalizáciu a emergenciu významu. Tento rámec zároveň neuzatvára diskusiu o povahe inteligencie, ale naopak, poskytuje nový impulz pre filozofické a epistemologické reflexie, ktoré nadväzujú na dlhú tradíciu uvažovania o vzťahu medzi výpočtom, poznaním a mysľou.

Záver

Historický vývoj umelej inteligencie ukazuje, že ide o oblasť, ktorá sa nevyvíjala lineárne, ale prostredníctvom postupných presunov dôrazu medzi rôznymi teoretickými a metodologickými prístupmi. Od raných úvah o efektívnych metódach a formalizácii myslenia, cez symbolické modely manipulácie so znalosťami, až po súčasné generatívne systémy možno sledovať

pretrvávajúcu snahu uchopiť inteligenciu ako proces, ktorý je v princípe modelovateľný pomocou výpočtových mechanizmov. Zároveň sa však ukazuje, že každý z týchto prístupov zvýrazňuje inú dimenziu inteligentného správania a prináša vlastné epistemologické obmedzenia. Symbolická umelá inteligencia zohrala v tomto vývoji kľúčovú historickú úlohu. Vytvorila prvý systematický rámec, v ktorom bolo možné explicitne pracovať s reprezentáciou poznania, logickým usudzovaním a riešením problémov. Jej sila spočívala v transparentnosti a interpretovateľnosti, no práve tieto vlastnosti sa ukázali ako nedostatočné pri pokuse zachytiť dynamický, situačný a skúsenostný charakter inteligencie. Kritika symbolizmu tak nepredstavovala jeho úplné odmietnutie, ale skôr poukázanie na hranice modelu, ktorý staval inteligenciu výlučne na manipulácii explicitných symbolov.

Nástup konektivistických a dátovo orientovaných prístupov znamenal zásadnú zmenu perspektívy. Inteligencia prestala byť chápaná ako aplikácia vopred definovaných pravidiel a začala byť interpretovaná ako emergentný jav vznikajúci z učenia, adaptácie a interakcie s prostredím. Tento posun umožnil praktické prekonanie viacerých problémov symbolickej AI, no zároveň priniesol nové otázky, najmä v súvislosti s interpretovateľnosťou, vysvetliteľnosťou a povahou reprezentácie významu. Súčasné generatívne modely tieto otázky ešte viac vyostrujú, keďže dokážu produkovať komplexné výstupy bez explicitnej symbolickej reprezentácie znalostí.

Z epistemologického hľadiska je však dôležité zdôrazniť, že generatívne modely nepredstavujú diskontinuitu vo vzťahu k základným princípom výpočtovej teórie. Fungujú v rámci rovnakého formálneho rámca, ktorý bol položený prácami Churcha a Turinga, a neprekračujú hranice Turingovej vypočítateľnosti. Zmena sa neodohráva na úrovni samotného výpočtu, ale v spôsobe, akým sú znalosti získavané, reprezentované a využívané. Tento fakt umožňuje chápať súčasný stav umelej inteligencie nie ako radikálne pretrhnutie s minulosťou, ale ako historicky konzistentné pokračovanie dlhodobého vývoja myšlienok o mechanizácii myslenia. Diskusia o umelej inteligencii sa tak nevyhnutne presúva od otázky „či“ je možné inteligenciu simulovať k otázke „akým spôsobom“ a „s akými dôsledkami“. Generatívne modely ukazujú, že vysoká úroveň funkčnej výkonnosti nemusí byť sprevádzaná tradične chápaným porozumením alebo intencionalitou. To otvára priestor pre nové filozofické interpretácie inteligencie, ktoré sa menej

sústreďujú na vnútorné mentálne stavy a viac na správanie, adaptáciu a schopnosť fungovať v komplexných informačných prostrediach.

Cieľom tohto článku nebolo poskytnúť definitívnu odpoveď na otázku povahy inteligencie, ale ukázať, že súčasné generatívne modely možno plnohodnotne pochopiť len v kontexte ich historického a konceptuálneho vývoja. Práve tento kontext umožňuje vyhnúť sa zjednodušeným interpretáciám, ktoré buď nekriticky glorifikujú súčasnú umelú inteligenciu, alebo ju naopak redukujú na čisto technický nástroj bez širšieho epistemického významu.

Záverom možno konštatovať, že vývoj umelej inteligencie predstavuje dlhodobý dialóg medzi formalizáciou a skúsenosťou, medzi explicitnými pravidlami a implicitným učením, medzi symbolom a štatistikou. Generatívne modely sú aktuálnym vyjadrením tohto dialógu, nie jeho uzavretím. Budúci výskum umelej inteligencie preto pravdepodobne nebude spočívať v nahradení jednej paradigmy druhou, ale v ďalšom hľadaní rovnováhy medzi rôznymi prístupmi k modelovaniu inteligentného správania a poznania.

Zoznam použitých zdrojov

BENDER, Emily M. et al. On the dangers of stochastic parrots. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency [online]. 2021, 610–623. DOI: 10.1145/3442188.3445922. [cit. 23. 11. 2025].

BUCHANAN, Bruce G.; SHORTLIFFE, Edward H. Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project [online]. Reading, MA: Addison-Wesley, 1984. ISBN 0-201-10172-6. Dostupné z: https://www.shortliffe.net/Buchanan-Shortliffe-1984/Chapter-01.pdf [cit. 12. 11. 2025].

COPELAND, B. Jack. The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life [online]. Oxford: Oxford University Press, 2004. ISBN 0-19-825079-7. Dostupné z: http://f.javier.io/public/books/The%20essential%20Turing.%20Seminal%20writings%20in%2

0computing,%20logic,%20philosophy%20-%20Copeland%20J.%20(2004)(ISBN%200198250800).pdf [cit. 23. 11. 2025].

DREYFUS, Hubert L. What Computers Can't Do [online]. New York: Harper & Row, 1972. ISBN 06-011082-1. Dostupné z: https://monoskop.org/images/c/ce/Dreyfus_Hubert_What_Computers_Cant_Do_A_Critique_of_Artificial_Reason.pdf [cit. 5. 11. 2025].

FLORIDI, Luciano. The Philosophy of Information [online]. Oxford: Oxford University Press, 2011. ISBN 978-0-19-923238-3. Dostupné z: https://www.academia.edu/27386529/FLORIDI_Luciano_The_philosophy_of_information_Oxford_Oxford_University_Press_2011 [cit. 12. 11. 2025].

GOODFELLOW, Ian et al. Generative adversarial nets. Advances in Neural Information Processing Systems [online]. 2014, 27. Dostupné z: https://papers.nips.cc/paper_files/paper/2014/file/f033ed80deb0234979a61f95710dbe25-Paper.pdf [cit. 12. 11. 2025].

HOBBES, Thomas. Leviathan [online]. Cambridge: Cambridge University Press, 1996. ISBN 521-56099-3. (Pôvodné vydanie 1651). Dostupné z: https://iranacademia.com/wp-content/uploads/2023/12/Thomas-Hobbes-Leviathan-1996.pdf [cit. 5. 11. 2025].

CHURCH, Alonzo. An unsolvable problem of elementary number theory. American Journal of Mathematics [online]. 1936, 58(2), 345–363. Dostupné z: https://ics.uci.edu/~lopes/teaching/inf212W12/readings/church.pdf [cit. 5. 11. 2025].

KLEENE, Stephen C. Introduction to Metamathematics [online]. Amsterdam: North-Holland, 1952. ISBN 0-7204-2103-9. Dostupné z: https://www.scribd.com/document/438963505/Kleene-Introduction-to-Metamathematics-1971-pdf [cit. 12. 11. 2025].

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. Nature [online]. 2015, 521, 436–444. DOI: 10.1038/nature14539. [cit. 23. 11. 2025].

McCARTHY, John. Recursive functions of symbolic expressions and their computation by machine, Part I. Communications of the ACM [online]. 1960, 3(4), 184–195. Dostupné z: https://www-formal.stanford.edu/jmc/recursive.pdf [cit. 5. 11. 2025].

McCARTHY, John; HAYES, Patrick J. Some philosophical problems from the standpoint of artificial intelligence. Machine Intelligence [online]. 1969, 4, 463–502. Dostupné z: https://www-formal.stanford.edu/jmc/mcchay69.pdf [cit. 12. 11. 2025].

McCARTHY, John; MINSKY, Marvin; SHANNON, Claude; ROCHESTER, Nathan. A proposal for the Dartmouth summer research project on artificial intelligence [online]. 1955. Dostupné z: http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf [cit. 23. 11. 2025].

McCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics [online]. 1943, 5, 115–133. Dostupné z: https://www.cs.cmu.edu/~epxing/Class/10715/reading/McCulloch.and.Pitts.pdf [cit. 23. 11. 2025].

MINSKY, Marvin; PAPERT, Seymour. Perceptrons [online]. Cambridge, MA: MIT Press, 1969. ISBN 0-262-63111-3. Dostupné z: https://rodsmith.nz/wp-content/uploads/Minsky-and-Papert-Perceptrons.pdf [cit. 12. 11. 2025].

MITCHELL, Tom M. Machine Learning [online]. New York: McGraw-Hill, 1997. ISBN 0-07-042807-7. Dostupné z: https://www.cs.cmu.edu/~tom/files/MachineLearningTomMitchell.pdf [cit. 5. 11. 2025].

NEWELL, Allen; SIMON, Herbert A. Computer science as empirical inquiry: Symbols and search. Communications of the ACM [online]. 1976, 19(3), 113–126. DOI: 10.1145/360018.360022. [cit. 5. 11. 2025].

NEWELL, Allen; SIMON, Herbert A. The logic theory machine. IRE Transactions on Information Theory [online]. 1956, 2(3), 61–79. DOI: 10.1109/TIT.1956.1056797. [cit. 23. 11. 2025].

RUSSELL, Stuart J.; NORVIG, Peter. Artificial Intelligence: A Modern Approach [online]. 4th ed. Harlow: Pearson, 2021. ISBN 978-1-292-40117-1. Dostupné z: http://lib.ysu.am/disciplines_bk/efdd4d1d4c2087fe1cbe03d9ced67f34.pdf [cit. 12. 11. 2025].

SEARLE, John R. Minds, brains, and programs. Behavioral and Brain Sciences [online]. 1980, 3(3), 417–457. DOI: 10.1017/S0140525X00005756. [cit. 23. 11. 2025].

TURING, Alan M. On computable numbers, with an application to the Entscheidungsproblem. Proceedings of the London Mathematical Society [online]. 1936, 42(2), 230–265. Dostupné z: https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf [cit. 23. 11. 2025].

VASWANI, Ashish et al. Attention is all you need. Advances in Neural Information Processing Systems [online]. 2017, 30. DOI: 10.48550/arXiv.1706.03762. [cit. 5. 11. 2025].

WINOGRAD, Terry; FLORES, Fernando. Understanding Computers and Cognition: A New Foundation for Design [online]. Norwood, NJ: Ablex, 1986. Dostupné z: https://www.academia.edu/2429378/T_Winograd_and_F_Flores_Understanding_Computers_and_Cognition_A_New_Foundation_for_Design [cit. 5. 11. 2025].

YATES, Frances A. The Art of Memory [online]. London: Routledge & Kegan Paul, 1999. ISBN 0-415-22046-7. Dostupné z: https://monoskop.org/images/b/be/Yates_Frances_A_The_Art_of_Memory.pdf [cit. 12. 11. 2025].

Autor

Mgr. Vladimír Filip, PhD.
vladimir.filip@umkd.uniza.sk

Ústav mediamatiky a kultúrneho dedičstva
Žilinská univerzita v Žiline

Univerzitná 8215/1
010 26 Žilina
SLOVENSKÁ REPUBLIKA

Mgr. Vladimír Filip, PhD. je absolventom doktorandského študijného programu Mediamatika a kultúrne dedičstvo na Katedre mediamatiky a kultúrneho dedičstva FHV UNIZA. Profesionálne sa zameriava najmä na životnosť digitálne zrodených dokumentov, pričom skúma nielen odolnosť a limity ich nosičov, ale aj problematiku dlhodobej archivácie, migrácie dát a záchrany ohrozeného digitálneho obsahu. Vo svojom výskume prepája technické aspekty digitálnej preservácie s metodologickými prístupmi digitálnych humanít a kultúrneho dedičstva.