

RICHARD SUSSKIND

**AKO PREMÝŠĽAŤ  
O UMELEJ  
INTELIGENCII**

PRÍRUČKA PRE  
ZMÄTENÝCH



# PRVÁ ČASŤ

Porozumieť AI

# PRVÁ KAPITOLA

## Leto AI

Ak chceme jasne uvažovať o AI, najprv si musíme zodpovedať otázku, čo je a čo nie je AI. Môj prístup k tejto veci je viac pragmatický ako teoretický.

Vo všeobecnosti sa dá umelá inteligencia definovať dvomi spôsobmi: podľa toho, ako funguje, alebo podľa toho, čo dokáže. Odborníci na technológie na ňu zväčša hľadajú prvým uvedeným spôsobom. Charakterizujú AI podľa techniky a technológií, na ktorých je postavená. Často pri tom používajú špecifický žargón. Hovoria napríklad o „neurónových sieťach“, „veľkých jazykových modeloch“, „spätnom šírení“, „grafických procesoroch“, „adversariálnom strojovom učení“ či „bázových modeloch“. Neraz tým ohromujú svojich poslucháčov. Väčšina ľudí, od strategických lídrov po bežných občanov, nepozná význam uvedených slov a nemá dostatočné technické vzdelanie na to, aby porozumeli konceptom, ktoré za nimi stoja.

Práve preto je mojím východiskom druhá zo spomenutých perspektív. AI sa dá podľa mňa najlepšie pochopiť prostredníctvom niekoľkých praktických príkladov toho, čo systémy s umelou inteligenciou dokážu už dnes a čo bude v ich silách už čoskoro.

### **Čo dokáže AI**

V čase počiatkov umelej inteligencie, v období takzvanej starej dobrej AI, sme umelú inteligenciu rozsekali na niekoľko od-

delených oblastí. Hovoríme o šesťdesiatych, sedemdesiatych a osemdesiatych rokoch minulého storočia. Používali sme vtedy pomerne všeobecné označenia a zameriavali sme sa na široké kategórie úloh pre AI. Jednou z nich bolo napríklad rozpoznávanie reči s cieľom vyvinúť systémy, ktoré by rozumeli slovám ľudí a dokázali by odlíšiť „nestiahni“ a „miesta, dni“. Ďalšou bol preklad z jedného jazyka do druhého tak, aby sme sa vyhli hrúbkam a nepreložili anglické „it’s raining cats and dogs“ (anglický výraz podobný nášmu „leje ako z krhly“; pozn. prekladateľa) ako „pršia mačky a psy“. Súvisiacou úlohou je „spracovanie prirodzeného jazyka“ umožňujúce ľuďom dávať počítačom pokyny a komunikovať s nimi ľudským jazykom, nie počítačovým kódom, a to so všetkými nepresnosťami a dvojnásobnosťami, ktoré to so sebou prináša.

Ďalšou vetvou AI sú expertné systémy. Najrozšírenejším prístupom je pokúsiť sa zachytiť a namodelovať procesy uvažovania a riešenia problémov špičkových expertov a vložiť ich do systémov. Takto by mohli aj laici dosahovať v mnohých sférach výkony porovnateľné so špičkovými expertmi.

Spomeňme aj ďalšie štyri oblasti, ktoré sa osvedčili ako základy nasledujúceho vývoja. Prvou sú hry, dodnes považované za úrodnú pôdu na testovanie AI a laboratórium pre pokrok najvyššej úrovne – nehovoríme len o šachu a čínskej hre go. Ďalšou oblasťou je robotika. Napriek náročným výzvam v poslednom čase výrazne napreduje. Treťou oblasťou je výskum zrakového vnímania. Dochádza k vývoju systémov, ktoré dokážu rozoznávať objekty (vrátane tváří) vo fyzickom svete. Napokon je tu automatické programovanie, čiže vývoj systémov, ktoré dokážu vyvinúť iné systémy. Môže ísť o najvýznamnejšiu oblasť, keďže v nej prechádzame od strojov vyvinutých výlučne ľuďmi k strojom, ktoré si dokážu napísať vlastný programový kód, samy sa vylepšujú a dokonca aj replikujú.

Prvotná éra „starej dobrej AI“ je za nami. Keď dnes hovoríme o tom, čo AI dokáže, zameriavame sa na konkrétnejšie úlohy, činnosti a oblasti. Mňa osobne najviac zaujíma využitie

AI pracovníkmi v odborných profesiách. Napríklad v oblasti zdravotníctva máme AI systémy, ktoré dokážu diagnostikovať choroby, interpretovať snímky, odporúčať liečbu, vykonávať robotické operácie, radiť pacientom a pomáhať pri vývoji liekov a vakcín (AI urýchlila vývoj mRNA vakcíny na covid od spoločnosti Moderna).<sup>9</sup> V oblasti práva poznáme systémy so schopnosťou navrhnúť, porovnať, zosumarizovať a zjednodušiť právne dokumenty a predpovedať rozhodnutia súdov.<sup>10</sup> Existujú daňové systémy, ktoré dokážu vytvárať a podávať daňové priznania a tvoriť plány a stratégie na odôvodnené znižovanie daňových povinností. V oblasti vzdelávania AI umožňuje personalizované učenie a prístup k najlepším výučbovým zdrojom na svete.<sup>11</sup> Navyše AI systémy dokážu navrhovať úžasné architektonické diela (a dokonca ich pomocou 3D tlače vedia aj skonštruovať).

Podobné pokroky zaznamenávame aj v oblasti manažérskeho poradenstva, auditu, prieskumov, poisťnej matematiky a v mnohých ďalších expertných a súvisiacich činnostiach, ako sú poisťovníctvo, bankovníctvo a investičný manažment. V prípade mnohých z týchto systémov sme ešte len na začiatku. Niektoré nie sú celkom spoľahlivé, hoci sa zdokonaľujú doslova zo dňa na deň. Iné sa už dnes vyrovnajú najlepším ľudským expertom. Ďalšie svojím výkonom už teraz predbehnú celé tímy najbystrejších špecialistov.<sup>12</sup>

Používanie AI nie je obmedzené iba na pracovníkov v odborných profesiách a zamestnancov v administratíve. Nájdete ju v mnohých sektoroch – od riešenia kvízov po riadenie elektrickej siete, od autonómnych vozidiel po kopilotovanie vojenských lietadiel, od optimalizácie obchodných procesov po plánovanie leteckých trás s cieľom znížiť emisie, od skúmania emocionálneho stavu človeka po vytváranie realistických hovoriacich tvárí. Umelá inteligencia sa čoraz viac dostáva do oblastí, kde bol potrebný výkon ľudského mozgu.

Mnohé AI systémy predstavujú samostatné produkty slúžiace na jediný účel (napríklad diagnostické systémy v zdravot-

níctve), no AI je čoraz častejšie zabudovaná v iných systémoch. Výrobcovia mobilných telefónov Apple a Samsung tvrdia, že ich prístroje sa dnes používajú jednoduchšie a lepšie, pretože sú vylepšené pomocou umelej inteligencie. AI sa integruje do softvérov, ktoré používame na každodennej báze. Microsoft ponúka svoj AI systém Copilot spolu so softvérom Office. Dokáže písať návrhy e-mailov a dokumentov, sumarizovať stretnutia a e-mailové konverzácie, vytvárať obsah prezentácií, zoznamy úloh a mnoho iného.

V praxi to znamená, že väčšina ľudí bude mať už čoskoro poruke vlastných osobných asistentov, ktorí im pomôžu pri získavaní informácií a organizovaní života. Budeme mať aj osobných učiteľov, ktorí nás bez váhania naučia to, čo budeme práve potrebovať. A to nie je všetko: neustále sa zdokonaľujúce AI systémy budú vedieť odpovedať na otázky z každej oblasti, dokážu vytvárať texty (výborne štylizované listy, poéziu, články, texty piesní), hudbu, kreslené rozprávky, kresby, krajinomalby, videá a veľa iného. AI dá takto ľuďom do rúk nové nástroje. Nejde iba o to, aby mohli kancelárski zamestnanci zvyšovať svoju produktivitu a efektivitu.

Už dnes používame umelú inteligenciu na každom kroku, no je pred nami skrytá. Keď nám obchodníci ponúkajú cez internet filmy, knihy, šaty či jedlo, ktoré by sa nám mohli páčiť, je za tým nejaká forma AI, voľne nazývaná algoritmus. Podobne AI beží niekde na pozadí, keď žiadame o zamestnanie, pôžičku alebo poisťku. Umelá inteligencia je v rôznych štádiách prítomná vo všetkých veľkých organizáciách, na ktoré sa spoliehame – od výroby po maloobchod, od dopravy po nemocnice, od bankového systému po štátnu správu.

Tieto systémy sa používajú aj na vytváranie nových svetov vo virtuálnej realite a dokonca aj na prepájanie mozgov s počítačmi (pozri jedenástu kapitolu). Niektorí nadšenci AI to dotiahli ešte ďalej a veria, že umelá inteligencia pomôže ľudstvu s najväčšími výzvami, pred ktorými stojí – od klimatickej zmeny po rakovinu, od chudoby po konflikty, od prieskumu

vesmíru po globálne vzdelávanie, od preľudnených miest po eróziu demokracie.<sup>13</sup>

Dúfam, že som povedal dosť na to, aby sme teraz mohli prísť s prvotným zhrnutím toho, čím je AI. Majte pritom na pamäti – a hovorí to dosť veľa o objekte nášho záujmu –, že ani samotní vedci zaoberajúci sa AI sa nedokážu zhodnúť na definícii umelej inteligencie. Keď dostalo 400 AI expertov úlohu vybrať si z navrhovaných definícií, najčastejšie vybranú z nich si zvolilo len 56 percent respondentov.<sup>14</sup>

Začal by som tým, že AI je odvetvie digitálnych technológií zamerané na vývoj a využívanie systémov schopných vykonávať úlohy a činnosti, o ktorých sa predtým predpokladalo, že ich môžu vykonávať iba ľudia prostredníctvom svojich pozoruhodných mozgov. Dnes musíme túto definíciu rozšíriť o tvrdenie, že AI systémy sú čoraz častejšie výkonnejšie ako ľudia a robia veci, ktoré sú mimo ľudských možností. (Niektorí komentátori a experti by chceli ísť ešte ďalej a vravia, že skutočná umelá inteligencia je iba tá, o ktorej môžeme povedať, že má „vedomie“. Z dôvodov, ktoré objasním neskôr, to však nepovažujem za potrebné a skôr sa nazdávam, že takáto podmienka môže vyvolať zmätok.)

Jedným z dôsledkov takto opísanej AI je, že otvára nové otázky v tradičných odvetviach. Sektor vzdelávania si musí dobre premyslieť, na čo má pripravovať študentov. Odborníci na investovanie, rizikový kapitál a súkromné investície sa musia zamýšľať nad tým, či tu ešte budú podniky, do ktorých investujú, keď sa rozhodnú speňažiť svoje výnosy. Experti na vojenskú techniku musia nanovo vymyslieť základné doktríny odstrašovania a obrany, pretože ich k tomu núti príchod autonómnych zbraní.<sup>15</sup> AI infraštruktúra národných štátov sa stáva potenciálnym zdrojom vojenskej výhody a hospodárskych rozdielov.

Pokroky umelej inteligencie by sme mali zohľadňovať aj v ďalších oblastiach, ktoré si vyžadujú dlhodobú perspektívu. Nie vždy sa to deje. Diskusie o klimatickej zmene, v ktorých sa

za mílnik často považuje rok 2050, zriedkavo zohľadňujú význam AI. Samozrejme, zaznievajú obavy o vysoké nároky AI na energetiku, ale ešte väčšou otázkou je škála zmien v našej civilizácii a ľudstve, ktoré vyvolá AI do roku 2050. V tejto knihe predpokladáme, že ľudstvo bude do roku 2050 zásadne iné ako dnes a AI systémy zásadne výkonnejšie ako ľudia. To nezníži intenzitu výziev súvisiacich s klimatickou zmenou, no nabáda nás to k tomu, aby sme sa na naše potenciálne potreby na rok 2050 pozerali zo širšej perspektívy.

Samozrejme, AI je pôvodcom závažných rizík a hrozieb. Venujem sa im v ôsmej a deviatej kapitole.

Nateraz sa teda zameriam na to, čo AI dokáže a čo by ešte dokázať mohla. Začneme tým, ako vlastne AI funguje. Najprv sa pokúsim bez použitia technického jazyka vysvetliť, ako prebiehal vývoj umelej inteligencie.

### **Kratučké dejiny AI**

Nazdávam sa, že v prvom rade je potrebné objasniť, prečo sa vyhýbam podrobným opisom fungovania AI systémov. Sčasti je vysvetlením cieľové publikum knihy – široká verejnosť bez nutnosti technického vzdelania. Hlavnou príčinou však sú rapidné zmeny, ktoré neustále prebiehajú v technike a technológiách AI systémov. Podľa môjho názoru by bol pohľad do útrov najnovších technológií z dlhodobého hľadiska pre väčšinu ľudí menej hodnotný ako pohľad na pravdepodobné možnosti týchto systémov v nadchádzajúcich rokoch. Do módy prídu nové systémy a metódy. Namiesto dôkladného skúmania podrobností dnešných technológií je užitočnejšie nazrieť na dejiny a trajektóriu vývoja umelej inteligencie. Prejdeme si preto rôzne mílniky v AI za ostatných 75 rokov. Ide o subjektívny pohľad – príbeh AI vyrozprávaný prostredníctvom úspechov, ktoré osobne považujem za najdôležitejšie.

Mnohí ľudia vedia niečo o AI zo science fiction. Významní autori tohto žánru – Isaac Asimov, Neal Stephenson, William Gibson, Philip K. Dick či Arthur C. Clarke – predpovedali mno-

hé dnešné i nadchádzajúce využitie umelej inteligencie.<sup>16</sup> Predpovedali aj súvisiace riziká. Spolu so scenáristami rôznych seriálov a filmov – najmä *2001: Vesmírna odysea* (1968), *Blade Runner* (1982), *Terminátor* (1984) či *Matrix* (1999) – títo autori výrazne sformulovali naše očakávania a obavy spojené s robotmi a so strojmi.

AI nenájdeme len v moderných filmoch, seriáloch a knihách. Názznaky umelej inteligencie môžeme vystopovať aj v starovekej klasike. Héfaistos mal v Homérovej *Iliade* dielňu, v ktorej vyrábal zbrane pre bohov. Boli tam „služobníci vytepaní zo zlata pripomínajúci živé služobníctvo“.<sup>17</sup> O viac ako 2 500 rokov neskôr sa aj v dielach z devätnásteho storočia vyskytujú pozoruhodné zmienky súvisiace s AI, napríklad vo *Frankensteinovi* (1818) od Mary Shelley či v románe *Erewhon* (1872) od Samuela Butlera. Aj začiatkom dvadsiateho storočia sa objavovali knihy s podobnou tematikou, napríklad diela H. G. Wellsa, zvlášť *Vojna svetov* (1895 – 1897). O pár desaťročí nato George Orwell v knihe *1984* (vydaná v roku 1949) pripravil pôdu na dnešnú debatu o všadeprítomnom elektronickom sledovaní občanov v spoločnosti.

AI prestala byť fikciou v roku 1950 po zverejnení prelomového článku Alana Turinga s názvom *Computing Machinery and Intelligence* (Výpočtová technika a inteligencia), v ktorom metodicky preskúmal možné odpovede na otázku, či stroje môžu myslieť (vraciam sa k tomu v desiatej kapitole).<sup>18</sup>

O pár rokov neskôr, v roku 1956, sa na Dartmouth College v americkom štáte New Hampshire stretla skupina vedcov na letnom workshope. V žiadosti o financovanie tejto udalosti bol prvýkrát použitý termín umelá inteligencia. Žiadateľom bol John McCarthy, vo všeobecnosti považovaný za autora termínu. Stretli sa tam iba talentovaní Američania.<sup>19</sup> Prvou medzinárodnou konferenciou o AI bolo sympóziu o „mechanizácii myšlienkových procesov“, ktoré sa uskutočnilo v Národnom fyzikálnom laboratóriu v anglickom Teddington v roku 1958.<sup>20</sup>

Prvý funkčný AI systém bol vyvinutý ešte pred vymysle-

ním termínu umelá inteligencia. Išlo o „Logic Theorist“, s ktorým v roku 1955 prišli a o rok nato ho dokončili informatik Allen Newell a ekonóm Herbert Simon. Inovatívnym prvkom tu bol vývoj systému presahujúceho tradičné činnosti, na ktoré sa v tom čase používali počítače: číselné operácie, aritmetika a dešifrovanie kódov. Ich systém sa vyznačoval nečíselným, symbolickým uvažovaním, ktoré v istom zmysle simulovalo ľudské premýšľanie. Objavil dôkazy pre teóremy v matematickej logike.<sup>21</sup> Práve tento systém opisuje Simon – ako sám hovorí, „s určitou pýchou“ – v citáte na začiatku tejto knihy.<sup>22</sup>

Ďalší mílnik sa opäť spája s touto prvotnou érou. John von Neumann urobil v sérii zapísaných prednášok, ktoré však už nestihol dokončiť a predniesť, prvú udržateľnú matematickú analýzu vzťahov medzi biologickým mozgom a počítačom.<sup>23</sup> Možno nepatrí medzi najznámejších vedcov svojej doby, no nepochybne ide o hviezdu matematiky a fyziky dvadsiateho storočia s nesmiernym vplyvom na vývoj počítačov a umelej inteligencie. Jeho uvažovanie je základom tejto knihy i mnohých ďalších súvisiacich kníh. Dnešné počítače sú stále konštruované na základe takzvanej von Neumannovej architektúry. V päťdesiatych rokoch pracoval na Princetone v Inštitúte pokročilých štúdií spolu s Albertom Einsteinom, Kurtom Gödelom či Robertom Oppenheimerom. A spomedzi nich práve jemu prischla prezývka Veľký muž.<sup>24</sup> Moja obľúbená historka o von Neumannovi – možno len vymyslená, ktovie, ale dobre som sa na nej pobavil – hovorí, že keď mal štyri či päť rokov, spýtal sa svojej mamy, ktorá fajčila cigaretu a uprene pri tom hľadela von oknom: „Mama, čo teraz počítaš?“<sup>25</sup>

Šesťdesiate roky priniesli viaceré prelomy vo vývoji AI, no dva spomedzi nich vynikajú. Prvým je iniciatíva Dendral, za ktorou stál Edward Feigenbaum zo Stanfordovej univerzity. Zachytil a v rôznych systémoch predstavil expertízu Joshuu Lederberga (nositeľa Nobelovej ceny za medicínu) a Carla Djerassiho (vynálezcu antikoncepčnej tabletky). Výsledný systém dokázal lepšie ako človek určiť štruktúru neznámych molekúl

na základe hmotnostných spektroskopických údajov. Dendral úspešne využil úplne nové techniky strojového učenia (pozri nižšie) a heuristiky (všeobecné pravidlá, ktoré v kombinácii s formálnymi poznatkami zabezpečujú výkon na najvyššej úrovni).

Druhým míľnikom toho obdobia bol pre mňa vývoj systému Eliza, za ktorým stál Joseph Weizenbaum, profesor AI na Massachusettskej technickej univerzite (MIT). Pomenovanie dostal po postave Elizy Doolittle z hry *Pygmalion* od Georgea Bernarda Shawa a autor ho vytvoril s istým nadhľadom. Napodobňoval v ňom interakciu medzi pacientom a psychoterapeutom. Weizenbaum požiadal svoju sekretárku, aby systém vyskúšala, a veľmi ho rozrušilo, keď ho vyzvala, aby odišiel z miestnosti, pretože ho chcela použiť v súkromí. Aj táto situácia ho podnietila do napísania jednej z najlepších kníh o vzťahu medzi ľuďmi a strojmi – *Computer Power and Human Reason* (Sila počítačov a ľudský rozum, 1984).<sup>26</sup> Predvídal v nej dnešnú debatu o etike AI a upozornil na riziká súvisiace s čoraz schopnejšími strojmi a ich vplyvom na ľudstvo.

Sedemdesiate roky boli poznačené úspechom Dendralu a z vedcov zameraných na AI sršal optimizmus. Snažili sa zmeniť teóriu na praktické aplikácie – prvotné „expertné systémy“. Jeden z nich, Prospector, pomáhal geológom pri vyhľadávaní ložísk nerastov. Vraj sa pomocou neho podarilo nájsť ložisko molybdénu za 100 miliónov dolárov. Najlepšie systémy z tohto obdobia vykonávali lekársku diagnostiku a odporúčali liečbu – Mycin sa zameriaval na infekčné choroby, Caduceus na internú medicínu a Casnet na zelený zákal. Všetky tieto systémy fungovali minimálne na expertnej úrovni.<sup>27</sup>

Napriek počiatočnému optimizmu a pokrokom zaznamenával vývoj AI v sedemdesiatych rokoch aj slabšie chvíľky. Na označenie obdobia, v ktorom výrazne upadá záujem o umelú inteligenciu aj investície do nej, používame termín zima AI. Skeptici takéto zimy milujú. Prvá nastala koncom sedemdesiatych rokov. Toto ochladenie prinieslo škrtý vo verejných výdav-

koch, najmä v USA, a názor, že dostupné technológie – hocako fascinujúce – zrejme nezabezpečia praktické výsledky a nenaplnia prísluby priekopníkov. AI nebola pre trh dosť presvedčivá.

S ústupom zimy však nastala začiatkom osemdesiatych rokov jar. Jej príchod bol spojený s nástupom „expertných systémov založených na pravidlách“, ktoré sa v tom čase považovali za aplikovanú, komerčne využiteľnú odnož AI. Podľa mňa sú expertné systémy prvou generáciou praktických AI systémov. Podnikateľská obec si veľmi rýchlo uvedomila, ako by ich mohla seriózne využiť v rôznych sektoroch vrátane práva, finančníctva, bankovníctva, poisťovníctva, maloobchodu, výroby a v mnohých ďalších.

Základom týchto systémov bolo vytváranie rozhodovacích stromov alebo vývojových diagramov, ktoré v širšom zmysle kopírovali uvažovanie expertov. Aj laickí používatelia dokázali pracovať s týmito rozhodovacími stromami. Systém kládol používateľom otázky, poskytoval odpovede a vytváral dokumentácie – opäť na expertnej úrovni.

Systémy osemdesiatych rokov boli mojím životným priestorom. V roku 1986 som na túto tému napísal dizertačnú prácu a potom, v rokoch 1987 až 1988, som spolu s popredným právnikom Phillipom Capperom vyvinul prvý plne funkčný komerčný AI systém pre právnikov na svete. Náš systém radil, kedy už nie je možné začať konanie pred súdom, pretože uplynulo priveľa času. V priebehu vývoja systému sme s Phillipom vytvorili obrovský rozhodovací strom s viac ako dvomi miliónmi ciest. Obsahoval všetky varianty okolností a právnych problémov, ktoré mohli nastať v tejto zložitej oblasti práva. Dá sa povedať, že sme zmenšili výklad príslušnej legislatívy a judikatúry na úroveň zložitého počítačového programu – znalosti sa dôkladne pretransformovali na formu kódu. Náš systém pomohol skrátiť čas vyhľadávania z hodín na minúty a každé rozhodnutie aj vysvetlil. Aj po viac ako 30 rokoch Phillip ochotne pripúšťa, že systém ho v mnohom prekonáva.<sup>28</sup>

Na prelome osemdesiatych a deväťdesiatych rokov prišla

druhá zima AI. Často sa spája s náročnosťou a vysokými nákladmi na industrializáciu expertných systémov a niekedy aj s kolapsom trhu špecializovaného AI hardvéru. Ja sa na to pozerám inak. Som presvedčený, že za tým stojí úsvit internetu. AI experti si na začiatku deväťdesiatych rokov všimli ľahkosť, s akou sa obsah a služby stávali dostupnými v online svete. Mnohí z nás, ktorí sme vtedy boli aktívni v oblasti AI pri podpore expertných činností, sa na pár rokov presunuli do vývoja webových nástrojov a aplikácií. Uznávali sme, že nie sú také sofistikované ako zložité expertné systémy založené na pravidlách, s ktorými sme narábali, no tvorba webových stránok v sebe mala nalievavosť, komerčný potenciál a pragmatizmus, vlastnosti, ktoré prilákali veľkú časť AI komunity.

Viacerí z nás sa až do polovice deväťdesiatych rokov umelou inteligenciou príliš nezaoberali. Keď však v roku 1997 majster sveta v šachu Garri Kasparov prehral s počítačom Deep Blue od IBM, boli sme v šoku. Niečo také by sme ešte v osemdesiatych rokoch nepredpokladali. Za jediný spôsob praktického využitia AI sme vtedy považovali zostavovanie rozsiahlych rozhodovacích stromov, ktoré zachytávali rozhodovanie expertov – presne tak, ako sme to urobili v oblasti práva. Najvýznamnejší experti, napríklad šachoví veľmajstri či lekári, však tvrdili, že v tých najnáročnejších situáciách, ktorým čelia, sa ich expertíza a skúsenosti nedajú zredukovať na vývojové diagramy. Boli presvedčení, že ich najlepšie ťahy či diagnózy podmieňovala akási nevysloviteľná, skrytá schopnosť – intuícia alebo inštinkt –, ktorú počítač nedokáže kódovať. Preto sme v osemdesiatych rokoch verili, že vždy bude existovať pevná deľba práce medzi strojmi a ľuďmi. Na rutinnú činnosť sa výborne hodia stroje, ale oblasť predstavivosti či tvorivosti je výsostnou doménou ľudí.

Mnohí tomu veria dodnes, no mýlia sa. Aj my sme sa mýlili v roku 1997, pretože sme nedokázali uchopiť exponenciálny nárast výpočtového výkonu počítačov (pozri druhú kapitolu). Keď Deep Blue porazil Kasparova, dokázal preanalyzovať viac

ako 300 miliónov možných ťahov za sekundu. Najlepší šachisti zvládnu zohľadniť približne 110 ťahov. Kasparova neporazil rovnocenný súper, ale hrubá výpočtová sila.<sup>29</sup>

Keď sme si mysleli, že stroje dokážu zvládnuť len rutinné činnosti, mýlili sme sa ešte v jednej veci. Objavil sa totiž nový spôsob vývoja systémov schopných dosiahnuť výkon ľudských expertov. Viedlo to k vzniku toho, čo nazývam druhou generáciou AI systémov.

Prvá generácia, expertné systémy, si vyžadovala „znanostných inžinierov“, ktorí dolovali cennosti z mozgov „doménových expertov“ a vytvárali z nich rozhodovacie stromy. Silnou stránkou týchto systémov bolo ľudské úsilie vynaložené na konštruovanie týchto vývojových diagramov. Samotné operačné systémy slúžili ako navigačné pomôcky. Na rozdiel od nich vykonávajú AI systémy druhej generácie veľkú časť náročných úloh. Nevyťahujú poznatky priamo od ľudí – dolovanie cenností prebieha v prostredí rozsiahlych dátových oceánov. Nemalo by nás prekvapiť, že všetky nazhromaždené digitálne údaje – či už sa nachádzajú na internete, v našich systémoch a databázach, alebo sú v textovom, prípadne multimediálnom formáte – zachytávajú, odzrkadľujú a obsahujú všestranné a kumulované skúsenosti, myšlienky, náhľady, expertné poznatky a emócie celého ľudstva. Technologický skok spočíval vo vývoji systémov, ktoré dokázali priamo odhaľovať vzory, trendy a vzťahy v týchto dátových oceánoch. Cieľom bol vývoj počítačových systémov, ktoré sa dokážu rôznymi spôsobmi „učiť“ z obrovských databáz bez toho, aby ich niekto musel explicitne programovať – tento jav označuje slovné spojenie strojové učenie. Takéto systémy sú „vycvičené“ na vykonávanie určitých druhov úloh, od diagnostiky chorôb po vypracúvanie návrhov zložitých dokumentov. Základnou technikou umožňujúcou strojové učenie sú „neurónové siete“. Veľkú časť priekopníckej práce v tomto smere vykonal nositeľ Nobelovej ceny Geoffrey Hinton v osemdesiatych rokoch, hoci aj jeho článok o „hlbokom učení“ z roku 2006 sa považuje za mimo-

riadny prelom v oblasti AI.<sup>30</sup> Neurónové siete – ako naznačuje už samotný názov – sú inšpirované štruktúrou ľudského mozgu. Dnešné systémy nereplikujú mozog. Sú to len „idealizácie“ toho, ako podľa nás funguje ľudský mozog.<sup>31</sup>

Prvý široko vnímaný úspech strojového učenia prišiel v roku 2011 a opäť za ním stála spoločnosť IBM. V živom vysielaní amerického televízneho kvízu *Jeopardy!* porazil jej počítač Watson (pomenovaný podľa zakladateľa IBM Thomasa J. Watsona) dvoch historicky najúspešnejších súťažiacich. Tento AI systém dokázal nájsť odpoveď na akúkoľvek otázku rýchlejšie a presnejšie ako ľubovoľný človek. Watson obsahoval prvky AI prvej aj druhej generácie – informácie v ňom boli explicitne kódované a zároveň sa opieral o strojové učenie z obrovského množstva údajov.<sup>32</sup> Nadšenci AI začali vnímať teplo nadchádzajúceho leta.

Potom, po pár rokoch, prišiel program AlphaGo, jeden z mojich obľúbených príkladov vysokovýkonného systému strojového učenia (hoci obsahuje aj prvky umelej inteligencie prvej generácie). Vyvinula ho spoločnosť DeepMind (v tom čase dcérska spoločnosť Googlu) na hranie zložitej hry go – takej zložitej, že umožňuje viac ťahov, ako je počet atómov vo vesmíre. Experti na AI nepredpokladali, že by niektorý AI systém dokázal hrať go na takej vysokej úrovni. Začiatkom roka 2016 však AlphaGo vyhral nad Leem Sedolom, špičkovým hráčom go, štyri ku jednej. Program využíval „hlboké neurónové siete“ a bol vycvičený prostredníctvom „učenia s učiteľom“ (využívajúc archív hier ľudských expertov) a „posilňovaného učenia“ (zdokonaľoval sa sám miliónmi partii odohraných proti samému sebe).<sup>33</sup>

Nadšenci AI často pripomínajú 37. ťah v druhej hre spomínaného zápasu z roku 2016. Program AlphaGo urobil ťah, na ktorý nikdy predtým neprišiel žiadny človek. Jeden z najlepších hráčov ho označil za prekrásny a povedal, že mu až vyhrkli slzy. Keby s takým ťahom prišiel človek, povedali by sme, že je kreatívny, inovatívny či dokonca geniálny. Nedá sa

to však povedať o ťahu stroja. Vznikol v dôsledku kombinácie masívneho výpočtového výkonu a pozoruhodných algoritmov fungujúcich na rozsiahlych dátach. Je zvlášť pozoruhodné, že niektoré ťahy programu AlphaGo úplne vybočovali z rámca uvažovania osôb, ktoré ho vyvinuli.

Niektorí odborníci namietali, že AlphaGo je síce mimoriadny program, ale jeho základom sú prevažne hry expertov z minulosti a nedarilo by sa mu tak dobre bez toho, aby čerpal z pôvodného ľudského talentu. Potom prišiel program AlphaGo Zero z roku 2017. Nenakrímili ho údajmi z historických hier. Naučili ho pravidlá a následne bol ponechaný sám na seba, aby si vypracoval optimálne stratégie a taktiku. AlphaGo Zero porazil pôvodný program AlphaGo vo všetkých sto hrách, ktoré medzi sebou hrali.<sup>34</sup> Verzia novšieho programu s názvom AlphaZero porazila majstrov sveta v šachu, šógi aj go.

V roku 2018 prišiel program AlphaFold od tej istej firmy. Podľa nositeľa Nobelovej ceny Venkiho Ramakrishnana išlo o ohromujúci pokrok v oblasti skladania bielkovín, 50 rokov starej obrovskej výzvy na poli biológie. K prelomu došlo o celé desaťročia skôr, ako odhadovali mnohí odborníci z branže.<sup>35</sup> Tento ich čin získal široké uznanie v roku 2024 – Demis Hassabis a John M. Jumper, dvaja AI vedci, ktorí viedli vývoj programu AlphaFold, dostali spoločne Nobelovu cenu za chémiu „za predpovedanie štruktúry proteínov“. Očakáva sa, že AlphaFold a jeho nástupcovia budú mať ďalekosiahly vplyv na liečenie chorôb a nakladanie s priemyselným odpadom.

### **Neskoršie prelomy**

Zatiaľ čo v DeepMind pracovali s prvou verziou AlphaFoldu, skupina AI vedcov z Googlu publikovala príspevok, ktorý priniesol ďalšiu významnú zmenu v pokroku umelej inteligencie.<sup>36</sup> Viedol k vzniku „transformátorov“, novej architektúry „hlbkového učenia“ pre neurónové siete. Na základe tejto architektúry sú postavené „velké jazykové modely“. Tieto systé-

my sú „trénované“ na obrovských databázach. Opäť sa nebudeme príliš zaoberať špecifickou terminológiou a fungovaním databáz. Zamyslime sa však nad slovami Stephena Wolframa, ktorý o tejto novej architektúre poznamenal, že „na prekvapenie všetkých dokáže neurónová sieť to, čo všetci považovali za jedinečnú ľudskú charakteristiku – tvoriť zmysluplný ľudský jazyk“.<sup>37</sup>

S využitím architektúry transformátorového modelu spoločnosť OpenAI, v tom čase americké výskumné laboratórium, uviedla na trh ChatGPT, dnes najznámejší veľký jazykový model. Musím povedať, že hoci som sledoval predchádzajúce verzie z radu GPT a dokonca som o nich aj písal, tento nový systém ma ohromil (GPT je skratka pre *generative pre-trained transformer*, čiže generatívny predtrénovaný transformátor). Išlo o najpozoruhodnejší prelom, aký som zažil počas svojich 40 rokov s AI. Spustenie tohto modelu znamenalo aj míľnik v tom, ako verejnosť vníma potenciál AI systémov. Zrazu sa zdalo, že máme k dispozícii nástroj schopný zodpovedať takmer každú zadanú otázku. Ešte úžasnejšie bolo, že keď sme ho vyzvali, aby napísal nejaký text – napríklad sonet, vyjadrenie sústrasti či dlhšiu esej na ľubovoľnú tému –, odpoveď prišla v rozmedzí sekúnd. Zdalo sa, že rozsah toho, čo dokáže, je neobmedzený.

Podobne ako mnohé obdobné systémy, ktoré prišli po ňom (napríklad Gemini od spoločnosti Google, Bing od Microsoftu či Claude od Anthropicu), je ChatGPT číťbot, čiže napodobňuje konverzáciu medzi ľuďmi. Označuje sa prívlastkom generatívny AI systém, čo znamená, že produkuje obsah na základe požiadavky. Nemusí ísť len o text – tieto systémy vedia vytvoriť skvelé umelecké obrázky, hudbu či videá. Poradia si dokonca aj s vysokokvalitným kódom – dokážu písať počítačové programy.

Pri systéme ChatGPT sa často zabúda na jednu dôležitú vec. Jeho prvá verzia predstavovala iba akýsi experiment. ChatGPT a generatívna AI nie sú významné tým, čím boli v roku

2022, alebo tým, čím sú dnes (ohromujú, no neraz sa dopúšťajú chýb), ale tým, že ich schopnosti sú predzvestou schopností budúcich systémov. Stále sme na začiatku. Treba si však uvedomiť, že tempo zmien a adaptácie sa zrýchľuje a že – ako si ukážeme v druhej kapitole – môžeme úplne zodpovedne predpokladať, že tieto systémy budú čoraz zdatnejšie a spoľahlivejšie.

ChatGPT a generatívna umelá inteligencia predstavujú najnovšiu kapitolu príbehu, ktorý siaha až do priekopníckych päťdesiatych rokov. Dnešné systémy nenahradili staršie expertné či prediktívne systémy ani nepredstavujú koniec príbehu. Umelá inteligencia je a aj zostane kumuláciou techník, pričom najaktuálnejšou z nich je generatívna AI.

AI je tu s nami už približne 75 rokov. V prvých dekádach jej vývoja prichádzali prelomy raz za päť až desať rokov. Dnes sa zdá, že tento čas sa skrátil na šesť mesiacov až rok; niektoré majú podobu technických inovácií, iné zásadných nových aplikácií. Keď v roku 2024 píšem tieto slová, nie je ľahké držať krok so zmenami v technológiách ani v praktickom rozsahu ich použitia. Nachádzame sa v horúcom lete AI. Podľa niektorých povedie dnešný rozmach a úroveň investícií k umelému a dočasnému prehriatiu, no nik pri zmysloch si nezakryje oči pred obrovskou úrovňou záujmu a investícií. Podľa mňa v dohľadnej budúcnosti slnko leta AI nezapadne.

Zdá sa, že umelá inteligencia sa dotkne všetkých sfér ľudského života. Kde sa to všetko skončí? O koľko pokročilejšími sa AI systémy ešte môžu stať? Sme blízko vytvorenia akejsi „superinteligencie“? Blížime sa k nejakému druhu „singularity“? Čo tieto termíny vôbec znamenajú? Objasním ich v druhej kapitole hneď po tom, ako zasadím AI do širšieho rámca digitálnych technológií.