

P R
O S T
O R

Henry Kissinger Craig Mundie Eric Schmidt Genesis

**Umělá intelligence,
naděje a lidská kreativita**

Přejeme si budoucnost, v níž se lidská a strojová intelligence navzájem posilují. Má-li nastat, musí jedna i druhá intelligence svému protějšku náležitě rozumět. Představa, že AI vycvičíme do té míry, že nám porozumí, a pak se prostě stáhneme a budeme doufat, že nás stroj začne respektovat, není zárukou úspěchu ani bezpečí. Musíme se snažit pochopit, před jaké výzvy nás AI postaví – být z dřívějška nemáme zkušenosti, které by nám zaručily, že situaci chápeme správně. A se zdoláváním tohoto nebetyčného úkolu musíme navíc zvládnout četné obtíže, před nimiž náš druh stojí již dnes.

P R
O S T
O R

**P R
O S T
O R**

**Henry
Kissinger
Craig Mundie
Eric Schmidt**

P R
O S T
O R

Henry
Kissinger
Craig Mundie
Eric Schmidt
Genesis

Umělá inteligence,
naděje a lidská kreativita

přeložil Martin Pokorný

PRAHA | 2025

GENESIS: ARTIFICIAL INTELLIGENCE, HOPE,
AND THE HUMAN SPIRIT

Copyright © 2024 by the Estate of Henry A. Kissinger,
Craig Mundie, and Eric Schmidt

Foreword copyright © 2024 by Niall Ferguson

First published in the United States by Little, Brown and Company,
a division of Hachette Book Group, Inc., 2024

All rights reserved

Czech edition © PROSTOR 2025

Translation © Martin Pokorný 2025

ISBN 978-80-7260-651-1

Knihu v úctě věnujeme
památce Henryho Kissinger:
státníka, diplomata, učitele a přítele.

Obsah

Předmluva – Niall Ferguson	13
Vzpomínka na Henryho Kissingera	25
Úvod	33
Část první: Na počátku	37
1. Objevitelství	39
2. Mozek	63
3. Realita	87
Část druhá: Čtyři odvětví	101
4. Politika	103
5. Bezpečnost	129
6. Prosperita	155
7. Věda	176
Část třetí: Strom života	193
8. Strategie	195
Závěr	223
Poděkování	229
Poznámky	233
Rejstřík	243
O autorech	252

Předmluva

Niall Ferguson

V červnu 2018 vydal Henry Kissinger v *The Atlantic* úvahu s názvem „Jak končí osvícenství“. Mnohé tehdy zarazilo, že tento kmet světové politiky, který krátce předtím oslavil pětadevadesátku, vůbec pocituje potřebu zaujímat k problému umělé inteligence nějaké stanovisko. Téma tou dobou ještě nebylo hitem jako dnes: vlna módního zájmu se vzedmula, teprve když firma OpenAI koncem roku 2022 spustila program ChatGPT.

Mým očím Kissingerova životopisce se však státníkův intenzivní zájem jevil naprosto přirozený. Vždyť Henry Kissinger původně dosáhl široké proslulosti publikací o jiné přelomové a zcela nové technologii. Kniha *Jaderné zbraně a zahraniční politika* (1957) vycházela

z výzkumu natolik důkladného, že jí vzdal hold sám Robert Oppenheimer: prohlásil ji za „mimořádně informovanou, a v tomto směru na poli jaderného zbrojení zcela bezprecedentní, (...) neboť s fakty zachází vysoce obezřetně a svou argumentaci, ač podávanou se zápallem, formuluje exaktně“.

Během doktorandského studia se sice Kissinger vnořil do dějin evropské diplomacie první poloviny 19. století, nicméně jinak si po celou profesní kariéru zřetelně uvědomoval, že odvěkou šachovnicí velmocenské politiky znovu a znovu otřásají proměny techniky. Shodně s mnoha svými vrstevníky, kteří se za druhé světové války zapojili do bojových akcí, viděl na vlastní oči nejenom hromadné ničení a zabíjení, jichž jsou moderní zbraně schopné, ale též osudné důsledky, jež pro Židy měla (jak to pamětihodně vyjádřil Churchill) „zvrácená věda“ Hitlerovy třetí říše.

Kissingera obestírala pověst válečného štváče. Nezaslouženě: ve skutečnosti ho po celý život vedl pocit povinnosti zabránit třetí světové válce coby nejobávanejšímu důsledku situace, v níž by studená válka mezi USA a SSSR přešla do horké fáze. Dobře si uvědomoval, že s využitím technologie jaderného štěpení by nový konflikt mezi velmocemi způsobil ještě větší zkázu než druhá světová válka. V *Jaderných zbraních a zahraniční politice* podává odhad ničivých důsledků, jaké by mělo shození bomby o síle 10 megatun TNT na New York – a z nich lze vyvodit, že kdyby Sověti plnou silou zaútočili na padesát největších amerických měst, přivodilo by to okamžitou smrt 15–20 milionů lidí. Dalších 20–25 milionů by utrpělo nevléčitelná těžká zranění. Pět až

deset milionů lidí by dříve či později zemřelo na důsledky radioaktivního spadu a přibližně 7–10 milionů dalších by neslo doživotní následky. Všichni přeživší by krom toho čelili „rozkladu společnosti“. Avšak i po takto pustošivém útoku, upozorňoval Kissinger, by Spojené státy zvládly stejně ničivě udeřit na Sovětský svaz. Z toho plynul zřejmý závěr: „Jediným výsledkem naplno vedené války by byla prohra obou účastníků konfliktu.“ Střet obdobného rozsahu nemá vítěze, argumentoval Kissinger ve stati „Strategie a organizace“ z roku 1957, „jelikož i slabší strana je schopna takové míry destrukce, kterou už žádná společnost neunesé“.

Idealismus však z mladého Kissingera neudělal pacifistu. V *Jaderných zbraních a zahraniční politice* zcela výslovně tvrdí, že „hrůzám jaderné války se nejspíš nevyhneme omezováním jaderné výzbroje“ ani inspekčními systémy. Otázka na pořadu doby tedy nezněla, zda se lze tomuto střetu zcela vyhnout, nýbrž zda „si dokážeme představit méně katastrofální uplatnění moci než naplno vedenou termonukleární válku“. Kdyby to totiž bylo vyloučeno, pak by Spojené státy se svými spojenci ve studené válce těžko zvítězily. „Absence jakýchkoli obecně uznávaných limitů válečného konfliktu,“ varoval Kissinger v článku „Kontroly, inspekce a omezená válka“ pro časopis *The Reporter*, „podlamuje psychologický rámec odporu vůči postupu komunistů. Zavládne-li mínění, že válka se rovná sebevraždě vlastního národa, může se kapitulace jevit jako menší zlo.“

O tento základ Kissinger opřel svou doktrínu omezené jaderné války, jak ji zformuloval ve „Strategii a organizaci“:

Na temně zlověstném pozadí termonukleární destrukce si už za válečný cíl nelze stanovovat vojenské vítězství v té formě, kterou mělo doposud. Cílem musí být dosažení určitých specifických politických podmínek, jež si protivník plně uvědomuje. Účelem omezené války je způsobit nepříteli ztráty nebo ho postavit před rizika zcela neúměrně překračující cíle, kvůli nimž střet vypukl. Čím je cíl uměřenější, tím vyšší je pravděpodobnost, že válka bude vedena méně násilně.

K tomu je ovšem nutné porozumět nejenom vojenským kapacitám, ale i psychice protivníka.

Kissingerova úvaha o omezené jaderné válce v daném období zdánlivým cynismem mnohé šokovala. Někteří experti – například Thomas Schelling – vyjádřili pochybnosti o tom, zda lze vůbec zamezit nekonečné eskalaci napětí, a od své argumentace se později distancoval i sám Kissinger. Přesto obě supervelmoci zkonstruovaly a využívaly bitevní či taktické jaderné zbraně v plném souladu s logikou, kterou Kissinger načrtl ve své knižní prvotině. Teoreticky byl snad koncept omezené jaderné války neobhajitelný, nicméně vojenští stratégové na obou stranách si počínali tak, jako by v praxi fungoval. (Příslušné zbraně existují dodnes a od chvíle, kdy se invaze na Ukrajinu zarazila v postupu, ruská vláda opakovaně hrozí jejich využitím.) Mladý Kissinger byl tedy v otázce jaderných zbraní blíž pravdě, než sám tušil.

O dopadech technologických změn v oblasti politiky ostatně autor uvažoval po celý život. V dnes již zcela přehlíženém materiálu, který v lednu 1968 vypracoval pro

Nelsona Rockefellera, předjímá dva způsoby, díky nimž by využití počítačů mohlo vládním úředníkům napomoci zpracovat stále mohutnější tok informací generovaných americkými státními agenturami. Podle Kissingerova mínění reálně hrozilo, že se vysocí představitelé začnou topit v záplavě dat. „Činitel ve vrcholné politice,“ podotkl k tomuto problému, „má k dispozici tolik informací, že se s nimi v krizových situacích vůbec nedovede vyrovnat.“ Všechny, kdo provádějí klíčová rozhodnutí, je třeba „soustavně seznamovat se situací v pravděpodobných centrech konfliktů“, zdůrazňoval Kissinger, a to i v těch případech, kdy příslušným ohniskům „nebyla přidělena vrcholná priorita“. Nadto je odpovědným činitelům nutné dodat „soubor praktických variant, (...) nastiňující hlavní alternativy v reakci na předvídatelné okolnosti, jakož i vyhodnocení pravděpodobných domácích i zahraničních důsledků každé takové volby“.

Kissinger si uvědomoval, že soustavné informační pokrytí by si vyžádalo zásadní investice do programování, utřídění, archivace i grafického zpracování dat. Naštěstí už ale podle něj existuje „technické řešení“ schopné zajistit všechny čtyři funkce:

Na jediný magnetický pás v délce 700 metrů dnes dokážeme uložit několik set osobních údajů o každém obyvateli USA. (...) Počítače třetí generace jsou nyní schopny provádět základní strojové operace v nanosekundách, to jest v miliardtinách sekundy. (...) Experimentální systémy sdílející výpočetní čas už prokázaly, že u velkých digitálních počítačů lze využívat víceuživatelský přístup umožňující vstup a výstup

informací na výkonných a konzolových stanicích distribuovaných po celém světě. (...) A zanedlouho budou pro počítačové výstupy k dispozici barevné katodové obrazovky.

Po prvním roce působení v Nixonově administrativě, v níž plnil funkci poradce pro otázky národní bezpečnosti, se Kissinger pokusil počítač tohoto typu získat pro osobní potřebu. CIA jeho žádost zamítla – nejspíš proto, že i bez něho zaměstnával rozvědné služby až dost.

Henry Kissinger se však z politiky nikdy tak zcela nestáhl a stejně tak neustal ani v úvahách o budoucnosti lidstva. Sotva se tudíž dalo předpokládat, že by s ohledem na svůj požehnaný věk přehlížel onen zásadní technologický zlom, k němuž došlo se vznikem a uplatněním generativní umělé inteligence. Naopak: úsilí porozumět dopadům rodící se technologie pro něj bylo v daném období jedním z hlavních okruhů zájmu.

Craig Mundie a Eric Schmidt – spoluautoři knihy, jíž se Kissingerova bibliografie uzavírá – jsou špičkoví technologičtí inovátoři a z textu sálá jejich bytostný optimismus. Očekávají „evoluci druhu *Homo technicus*“ – tedy člověka schopného žít „s mechanickou technologií v symbióze“. Síly umělé inteligence bude podle nich brzy možné „zapráhnout tak, aby vygenerovaly nový standard lidského bohatství i blahobytu“ a přinesly „odlehčení, ne-li přímo naprostou eliminaci pracovní a třídní zátěže – a spolu s tím i břemene konfliktů, které lidstvem zmítají“. Soudí dokonce, že přechod na AI by mohl přinést zásadní vyrovnání životní úrovně „bez ohledu na rasu, pohlaví, národnost, místo narození a rodinný původ“.

Příspěvek nejstaršího z trojice spoluautorů je nicméně zřetelně patrný v sérii varování, jež se z této práce vynořují jako leitmotiv. „S nástupem umělé inteligence,“ konstatuje kniha, „je v sázce samo přežití lidstva.“ AI, která nepodléhá náležité kontrole, „by mohla hromadit znalosti za účelem destrukce“. Snadno v této formulaci rozeznáme otázku nastolenou již v Kissingerově článku „Jak končí osvícenství“ z roku 2018:

Objektivní schopnost [umělé inteligence] dosahovat ne-lidskými metodami nových a správných poznatků o našem světě otřásá naši loajalitou k vědecké metodice, soustavně prosazované posledních pět staletí, a zpochybňuje též nárok lidstva na to, že výhradně ono rozumí povaze skutečnosti. Co to ve výsledku znamená? Máme se obávat, že namísto toho, aby éra AI přinesla lidstvu pokrok, spustí návrat k předmodernímu uctívání autorit? Hrozí, že se tím ocitneme na pokraji ohromného zvratu v lidském poznání – na prahu temného osvícenství?

V oddíle, který na mne osobně zapůsobil nejpronikavěji, se autoři zamýšlejí nad hluboce zneklidňující představou závodů ve zbrojení s využitím AI. „Pokud každá lidská společnost (...) touží jednostranně utužit vlastní převahu,“ píší, „pak se tím připravuje bitevní pole pro psychologický souboj konkurenčních vojsk a výzvědných služeb v míře, jaké lidstvo doposud nikdy nečelilo. V letech, měsících, týdnech i dnech vedoucích ke zrodu první superinteligence se před námi rýsuje bezpečnostní dilema existenčního významu.“

A pokud soutěží „o vytvoření jediné, dokonalé a nepochybnitelně dominantní inteligence“ přihlížíme již dnes, jaké výsledky se tedy rýsují? Autoři načrtávají šestici scénářů, přičemž ani jeden není lákavý:

1. Při soutěžení o existenci mezi množstvím aktérů, kteří stanuli před bezpečnostním dilematem, přijde lidstvo o výsadu kontroly.
2. Lidstvo přemožené vítězem, jehož nepoutají brzdy ani rovnovážné stavy historicky potřebné k tomu, aby ostatním zůstala záruka minimální bezpečnosti, ztratí svrchovanou hegemonii.
3. Ve světě se neprosadí jedna jediná svrchovaná AI, nýbrž vynoří se několik různých podob mocnější inteligence.
4. Firmy, které vlastní a rozvíjejí umělou inteligenci, na sebe strhnou společenskou, hospodářskou, vojenskou a politickou moc, díky níž se jim ocitne nadosah totální nadvláda.
5. Umělá inteligence se projeví nejdouhodobějším i nejvšudypřítomnějším způsobem a nejvyššího významu dosáhne v náboženských – nikoli státních – strukturách.
6. Nekontrolované šíření nové technologie formou otevřeného zdrojového kódu může způsobit vznik menších skupin či kmenů disponujících sice podřadnou, a přesto značnou AI kapacitou.

Těchto a podobných scénářů se Kissinger upřímně obával – a jeho snaha se jim jakkoli vyhnout sepsáním této knihy ještě neskončila. Není tajemstvím, že posledním projektem, do něhož těsně před svými stými

narozeninami vléval zbývající energii a kterému se oddal právě ve snaze tyto dystopické hrozby odvrátit, bylo iniciovat rozhovory o omezení zbrojení v oblasti AI mezi Spojenými státy a Čínou.

Závěr Genesis nese jeho nezaměnitelný rukopis:

Co je pro jedny kotvou, která nás drží vprostřed bouře, je pro druhé zdou, jež nám brání v pohybu. Co jedni vychvalují jako nezbytné kroky k dosažení vrcholku lidského potenciálu, jeví se druhým jako svévolný pád do propasti.

Instinktivní emocionální rozdíly – a subjektivní hranice, které vytyčí každá ze zúčastněných stran – tu vytvoří nepředvídatelnou a výbušnou situaci. Tlak okolností se dále umocní pod vlivem toho, jak se rozevřou nůžky mezi osudem „vítězů“ a „poražených“. Bojácní budou brzdit vlastní vývoj a sabotovat vývoj u druhých. Sebejistí se budou snažit maskovat vlastní kapacity a potajmu práci urychlovat. Časová osa nadcházejících krizí nabere zcela bezprecedentní tempo, které nás brzy zavalí, a není vůbec jisté, zda a jak přežijeme.

Technicky založený člověk na tak zlověstné předtuchy nejspíš odpoví, že AI lidstvu přináší hmatatelný prospěch, již dnes nepřehlédnutelný v oboru medicíny. S tím se dá jediné souhlasit. AlphaFold – model založený na neuronové síti a předpovídající trojrozměrné struktury proteinů – představoval podle mého soudu mnohem zásadnější zlom než ChatGPT. Nicméně srovnatelného pokroku medicína dosahovala i ve 20. století, a přesto – ruku v ruce s objevy a širokou dosažitelností

antibiotik, nových vakcín i bezpočtu dalších léčebných prostředků – v něm došlo ke dvěma světovým válkám a k hrůzám holokaustu.

Koneckonců zásadní problém technického pokroku jasně vyvstal v období, jehož byl Kissinger očitým svědkem. Jaderné štěpení objevili roku 1938 v Berlíně němečtí chemikové Otto Hahn a Fritz Strassmann. Teoretické objasnění fenoménu podala roku 1939 dvojice fyziků původem z Rakouska – Lise Meitnerová a její synovec Otto Robert Frisch. Na proveditelnost řetězové jaderné reakce, která by vedla k „výrobě energie a radioaktivních prvků, a bohužel možná rovněž k atomovým bombám“, poukázal maďarský fyzik Leó Szilárd. A zároveň odborníci pochopili, že by se tato řetězová reakce dala využít v jaderném reaktoru pro výrobu tepla – avšak zatímco sestavit první atomovou bombu trvalo vědcům jen něco málo přes pět let, první jaderná elektrárna zahájila provoz až roku 1951.

Položme si nyní otázku, čeho lidé za posledních osmdesát let zkonstruovali víc: jaderných hlavic, či jaderných elektráren? Ve světě se dnes nashromáždilo přibližně dvánáct a půl tisíce jaderných hlavic a jejich počet stoupá v důsledku toho, s jakou prudkostí Čína navyšuje svůj jaderný arzenál. V porovnání s tím je v provozu 436 jaderných reaktorů. V absolutních číslech dosáhla jaderná energetika vrcholu roku 2006 a její podíl na celosvětové výrobě elektřiny klesl z 15,5 % v roce 1996 na 8,6 % v roce 2022 – zčásti v důsledku přehnaných politických reakcí na několik havárií, jejichž dopad na lidské zdraví a životní prostředí je přitom ve srovnání s dopadem emisí oxidu uhličitého z fosilních paliv vcelku zanedbatelný.

Doba Henryho Kissingera nám udílí jasně čitelnou lekci: technologický pokrok může prospívat i ničit – v závislosti na tom, jak se ho rozhodneme kolektivně využít. Umělá inteligence a štěpení jádra se pochopitelně v mnoha směrech liší. Bylo by však zásadní chybou uvěřit domněnce, že tuto technologii zcela jistě využijeme ve prospěch tvořivých cílů namísto k ničení.

Historická i osobní zkušenost Kissingera poučily o tom, jak důležité je mít toto dilema trvale na paměti. Právě proto svůj život podstatnou měrou zasvětil promýšlení světového řádu a snaze zabránit další světové válce. Proto na nedávné průlomy v oblasti umělé inteligence reagoval tak hbitě a s takovými obavami. A proto je tato posmrtná publikace důležitá stejně jako všechny práce, které během svého dlouhého a vlivného života vydal.

Oxford, červenec 2024

Vzpomínka na Henryho Kissingera

Henry A. Kissinger zesnul 29. listopadu 2023. I ve svých sto letech představoval inspirativní vzor pro všechny, kdo ho kdy dostali příležitost poznat, a na této knize – své dvaadvacáté – pracoval do posledního dechu. Při našich častých schůzkách v posledním roce života odhodlaně zdůrazňoval, o jak závažné téma jde a jak naléhavě nezbytné je seznámit s ním širokou veřejnost. Jakožto spoluautoři jsme se ocitli v úzké skupince posledních, kdo s ním krátce před smrtí mluvili. Když se nyní celý projekt – jeho jménem a na jeho žádost – realizuje, je naším cílem uchovat originalitu jeho myšlení i zřetelný tón jeho hlasu v otázce, která je pro budoucnost lidstva vrcholně důležitá. Snahu dokončit, co započal – tak,

aby jeho poslední tvůrčí projekt nezemřel spolu s ním, nýbrž žil dál ve světě, z něhož sám musel odejít – chápeme jako skromný hold jeho památce.

Poté, co se tak významně zasadil o vybudování našeho současného světa, vynaložil veškerý čas až do svých posledních dnů a hodin na jeho záchranu. Jeho posledním písemným vyjádřením je výzva, abychom neustali v úsilí o zajištění budoucnosti lidského druhu. V polovině 20. století byl hlavním architektem filozofické a diplomatické snahy zachránit lidstvo před atomovou katastrofou, moderní podobou extrémního existenčního rizika. Tento odvážný obránce před někdejší hrozbou nás však opustil právě ve chvíli, kdy se před našimi zraky zjevuje další. Jeho žití se uzavřelo v době, kdy se rodí nová podoba života. Dnes, na úsvitu éry umělé inteligence, si palčivě uvědomujeme, jak mimořádné předpoklady měl – coby znalec devatenáctého, učitel dvacátého a prognostik jednadvacátého století – pro to, aby položil spolehlivé základy naší budoucnosti.

Byl v prvé řadě filozofem dějin. Přes pronikavé úvahy na téma tragiky existence dospěl k celoživotní snaze předvést, že zušlechťující idealismus srdce je slučitelný s realismem mozku. Řečeno slovy francouzského spisovatele Romaina Rollanda: „pesimismus intelektu“ lze zastávat společně s „optimismem vůle“. Zatímco optimista aspiruje na to, aby jednotlivec získal kontrolu nad vlastní existencí, pesimista vnímá, že na nás doléhá tlak přírodních zákonů a historických cyklů – sil, jež se naši nadvládě vymykají.

Henry Kissinger si dobře uvědomoval, že se idealistického zápalu mohou zmocnit ideologové schopní

nemilosrdně prolévat krev a tvrdě prosazovat moc. Nacionalismus, komunismus, totalitarismus i militantní náboženský fanatismus vznášely nároky na ryze idealistické cíle. Kissinger se zprvu coby obět těchto nelidských excesů a poté jako jejich vojenský a diplomatický protivník rozhodl pomoci vybudovat nový svět na principu sebevědomého řádu a bezpečnosti bez výmluv. Po celé období, v němž se aktivně podílel na řízení mezinárodních vztahů, ukazoval své adoptivní vlasti i celému světu cestu skrz bouře nejistot, opřen o pevnou půdu historických faktů i národního zájmu.

Jakkoli byl brilantním obhájcem selektivní nezbytnosti realismu, zůstával současně idealistou, který (řeceno slovy Nialla Fergusona, autora jeho životopisu) respektuje „úlohu lidské svobody, volby a aktivního jednání při formování světa“. Teoretickými úvahami i praktickými kroky manifestoval přesvědčení, že lidé nežijí a ani nemohou žít s pocitem, že budoucnost je nevyhnutelně předurčena. Již v diplomové práci o smyslu dějin se jakožto sedmadvacetiletý student Harvardu potýkal se shodnými filozofickými problémy, jimiž je posléze prodchnuta i jeho poslední kniha: „Nezávisle na tom, jak přesně chápeme nutnost dění kolem nás [a] (...) jak přesně ve zpětném pohledu vysvětlujeme lidské jednání, je jedna věc jistá: konatel byl při nich vnitřně přesvědčen, že se samostatně rozhoduje.“

Zda lidstvo skutečně přežije nástup nelidských zařízení, ukutých v žáru našich vlastních výhní, to Kissinger vůbec nepovažoval za jisté. Ve střetu s tíživou hrozbou jaderné katastrofy nepodlehł deterministickému fatalismu ani prorokům zkázy. Obavy o existenci sice mohou

zavdat podnět k nihilismu, ale stejně tak mohou dobré lidi prodchnout energičností nezbytnou k tomu, aby se zlu vzepřeli a uhájili to, co je pro budoucnost našeho druhu nepostradatelné. Na počátku padesátých let se Kissinger coby mladý akademik na Harvardu zúčastnil série schůzek, na nichž přední specialisté diskutovali o možných důsledcích jaderné války a o tom, jaká preventivní opatření je nutno přijmout. Plodem těchto schůzek byly zásady, které od té doby úspěšně chrání náš svět před nejhoršími obavami tehdejších účastníků.

V rozhovorech s námi se k těmto sedmdesát let starým schůzkám – k jejich struktuře, účelu i dnes již očividnému historickému významu – opakovaně vracel. A svůj přístup si uchoval do poslední chvíle: odmítal podléhat tlaku osudu, avšak neoddával se ani utopickým vizím. V oblasti umělé inteligence vyvstává nutnost téže rovnováhy jako v případě jaderného štěpení: hrstka věci oddaných jednotlivců může změnit chod dějin tím, že do nich zasáhne a dá najevo, jaké vyznává hodnoty. Zároveň byl ale Kissinger bez ohledu na genialitu vědců, kteří tuto novou inteligenci vyvíjejí, přesvědčen, že jejich školení samo o sobě nepostačuje k tomu, aby se při zacházení s nejnovějšími nástroji zachovalo nutné minimum zabezpečení.

Jeho odkaz v oblasti AI se proto neomezuje na čistě filozofické a teoretické úvahy. Zahrnuje i praktické návrhy. Půl století od tajné návštěvy Pekingu a následného navázání vztahů mezi Spojenými státy americkými a Čínskou lidovou republikou podnikl další přelomovou cestu do čínské metropole, když se na naléhavé a jmenovité pozvání prezidenta Si Ťin-pchinga vypravil

projednat rizika, jimž lidstvo v souvislosti s AI čelí. Byla to jeho vůbec poslední zahraniční cesta a poslední diplomatická mise.

Jestliže v dobách minulých pozvedl znalost a praxi diplomacie na úroveň umění, dnes svým pátráním po odpovědích pozvedl umělou inteligenci v cosi, co překračuje hranice čisté vědy. Spolu s jedním z níže podepsaných a s prof. Danielem Huttenlocherem z MIT vydal v roce 2021 knihu *Věk umělé inteligence a naše lidská budoucnost*, zahrnující předpověď, že nástup umělé inteligence vytyčí novou éru vyznačující se podobně zásadními a nečekanými změnami v lidském myšlení, jaké v 18. století přinesl nástup osvícenství. Avšak namísto aby lidstvo vycházelo z otázek kladených lidmi, konfrontuje se v této nové éře s odpověďmi umělé inteligence na otázky, které jí žádný člověk nepoložil. Úměrně tomu, jak umělá inteligence dobývá další a další oblasti lidského poznání, se Kissinger snažil čerpat z hlubin zdrojů lidské moudrosti.

V této knize společně s ním zkoumáme, jaký dopad má umělá inteligence na osm různých domén lidské aktivity a myšlení – načež celý výklad vyústí do jeho vlastní filozofické odpovědi na trvalé hledání schůdné strategie při vyvažování zisků a rizik. Za tímto účelem Kissinger zkoumá vyhlídky na koexistenci člověka s umělou inteligencí i na souběžný vývoj člověka a AI. Konceptuálním zpřístupněním možnosti smíru těchto dvou forem života, organické a syntetické, rovněž odkrývá nutnost rozhodnutí: buď vytvoříme svět, v němž se AI přibližuje nám, anebo svět, v němž se my sami připodobňujeme AI.

Počínaje rokem 2021, kdy o éře AI vydal svou první knihu, si Kissinger stále zřetelněji uvědomoval, že užitečnost rozumu v posledních dnech této éry naráží na své meze. Objasnění, která se vymykají našemu chápání nebo k nimž jsme nedospěli sami, se nám lidem mohou zdát naprosto šokující – přičemž instinkt nám velí předpokládat, že jsou oproti našim vědeckým výkladům čímsi zaostalejším a primitivnějším, tedy krokem zpět, a ne kupředu. Jenže to je nebezpečná domněnka.

Pokud je (jak napsal Arthur C. Clarke) „každá dostatečně rozvinutá technologie k nerozeznání od kouzelnictví“ a pokud sestrojování zázraků vychází z matematických principů, pak by budoucnost měla být neobjasnitelná, záračející či přímo magická. Po desítky let, kdy jsme působení Henryho Kissingera sledovali, nás svou obeznámeností s uměním státnickým – s touto složitou říší lidských záležitostí – velkoryse učil, že ačkoli čistá racionalita vskutku představuje ono hlavní paradigma, díky němuž lidé opanovali svět, nemůže být vzorcem, jehož prostřednictvím získáme vládu sami nad sebou.

Jakkoli byl tedy rozum historicky vzato motorem našich největších úspěchů, neměli bychom očekávat, že na něj lze i v budoucnu plně spoléhat. Zároveň se ho ale nemůžeme jednoduše zříct. Obdobou umírněného stanoviska mezi idealismem a realismem je i to, jak Kissingerovy poslední průzkumy budoucnosti lidstva dosahují rovnováhy mezi empiričností pravdy a čímsi jiným, co se rozumu vymyká, nicméně časově po něm následuje. Stejně jako si rozhodování na poli zahraniční politiky nemůže dovolit výstřelky v jednom ani ve druhém směru, musí se krajnostem vyhnout i rámec naší budoucnosti.

Umělá inteligence je tedy bezprecedentní výzvou a žádá si zamyšlení, které se na první pohled může zdát iracionální nebo přepjaté – a scénáře načrtnuté v této knize jsou nepochybně zářející. Když nás ale Henry Kissinger upozorňoval, že on sám je pouze skromným analytikem lidské povahy a jejího nejnovějšího, ne-li snad posledního výtvoru, současně nám naléhavě vštěpoval, že největším rizikem, před které nás umělá inteligence staví, by bylo unáhlené či příliš kategorické prohlašování, že jí rozumíme.

S pronikavostí jeho intelektu a vnímavostí k lidem se sotvakdy znovu setkáme. Nevíme o nikom jiném, kdo by byl ještě po devadesátce schopen si do detailů osvojit naprosto novou, neprobádanou doménu odborných znalostí. Neutišitelná zvědavost a bystrost se u něho družily s pracovitostí a cílevědomostí, a proto v něm žádné tělesné ani duševní trápení nedokázalo utlumit objevitelský a myslitelský zápal. Přes nevyhnutelné obtíže stáří projevoval den co den neochvějně odhodlání posunout svět kupředu. Snad tuto nevyčerpatelnou vnitřní sílu čerpal z maximální ukázněnosti, kterou v něm v mládí utužil útlak a posléze upevnila jak služba v armádě, tak desítky let tvrdých zkoušek v aréně veřejných diskusí a sporů.

Tento mimořádný muž zapůsobil na mnohé a my nejsme tak troufalí, abychom si mezi nimi osobovali zvláštní místo. Bude nám scházet – v danou chvíli ještě sami sotva tušíme, jak hluboce. Naše doba stojí na prahu ohromných nejistot a lidí jako on je dnes třeba více než kdy jindy. Proto se nám pro knihu zdál vhodnou volbou název *Genesis*: nový začátek pro něj i pro celé

lidstvo. Ať už bude lidstvo ve svém konání úspěšné, či ne, on sám již konečné důsledky svých snah nespatří. Přinejmenším se ale můžeme řídit jeho uvážlivostí a myšlenkovou pronikavostí.

Eric Schmidt a Craig Mundie

Úvod

Ještě před několika lety se umělá inteligence (AI) krčila na okraji veřejného zájmu. V mezidobí však hned v několika ohledech došlo k zásadnímu průlomů, téma umělé inteligence se ocitlo na titulních stránkách novin a touto problematikou se nyní zabývají vůdčí osobnosti v oblasti vědy, podnikání, žurnalistiky, veřejné administrativy, školství i politiky na celém světě.

Široké veřejnosti i řadě odborníků ale podle našeho mínění ještě dnes unikají četné klíčové aspekty nové éry. Příští podoby umělé inteligence a způsob, jak na ně lidé zareagují, mohou bytostně proměnit vztah lidstva ke skutečnosti a k pravdě, mohou transformovat způsob získávání poznatků i náš tělesný vývoj, ovlivnit

diplomatická jednání i mezinárodní uspořádání. Pro nejbližší desetiletí se jedná o zásadní otázky, o něž by se měli zajímat reprezentanti všech oborů.

Přestože nám nynější schopnosti umělé inteligence připadají obdivuhodné, zpětně se budou zdát chabé: kapacita AI se rozvíjí stále rychleji a naši všední existenci prostoupí možnosti, jaké si v danou chvíli ani nedokážeme představit. Budoucí systémy podnítí ohromný a převážně pozitivní pokrok, vylepší náš zdravotní stav a vygenerují bohatství.

Zároveň však s sebou tento rozvoj nese technická a lidská rizika: zčásti již známá, zčásti netušená. Již dnes nové technologie fungují způsobem, který jejich vynálezci nedokázali předvídat, a totéž lze očekávat i do budoucna. Každá produktivní linie výzkumu, po níž se odborníci vydají – a takových linií se rýsuje vícero –, může přinést nové domény nepředvídaných schopností, u nichž není předem zaručeno, zda je člověk dokáže pochopit a zda mu z nich plyne prospěch.

Pod vlivem AI se zdánlivě smrskávají časová měřítka, takže se budoucnost jeví mnohem blíž než dřív. Stačí uvést jeden příklad: blíží se zrod strojů schopných stanovovat si vlastní cíle. Pokud chceme vůbec doufat, že s nadcházejícími riziky udržíme krok, je třeba reagovat co nejrychleji. S vědomím toho, co je v sázce a jak naléhavé úkoly před námi stojí, zde upozorníme jen na několik fazet onoho problému.

S tím, jak se partnerství člověka a stroje stává všudypřítomným, se budou lidé muset rozhodnout, jaká je vlastně povaha těchto vztahů. Odpověď může vzejít z logiky bezpečnosti a efektivity, může ji poskytnout

pohled do dějin nebo nám ji může vnuknout cosi božského. Jedinci, národy, kultury a náboženství budou muset vytyčit, kudy – pokud vůbec někudy – vedou meze, kterým autorita AI nad pravdou podléhá. Budou muset rozhodnout, zda umělé inteligenci povolí roli prostředníka mezi lidmi a skutečností, nicméně v této souvislosti se ocitnou před dilematem, zda zachovat tradiční úlohu lidské iniciativy (a přitom nejspíš umělé inteligenci předat kormidlo při objevování nových poznatků), anebo se biologicky omezené lidské mysli zříct ve prospěch nově koncipovaného partnerství s AI na intelektuálním předvoji. Zvolíme si vlastní cíle a zapojíme AI do procesu jejich dosahování? Nebo povolíme, aby některé cíle určila sama? Nejnaléhavější je však požadavek, aby lidstvo zformulovalo moderní a udržitelnou definici lidské důstojnosti, podle níž se v nadcházejících letech budeme moci filozoficky při rozhodování orientovat.

S nástupem umělé inteligence je dle našeho názoru v sázce samo přežití lidstva. Jak vysvětlíme níže, budoucí kapacity AI – operující nadlidskými rychlostmi – zúčtují s tradičními formami regulace a lidé budou potřebovat fundamentálně nové kontrolní mechanismy. Aktuálním úkolem, před nímž globální vědecké společenství stojí, je vyvinout technická opatření pro to, abychom do všech AI systémů mohli zabudovat vnitřní pojistky. A až jednotlivé státy i mezinárodní organizace dosáhnou konsenzu, budou muset vypracovat nové politické struktury za účelem monitoringu, prosazování pravidel a reakcí na krizové situace. To si vyžádá vyřešení ne jednoho, nýbrž dvojího „vyladění kurzu“ – totiž technického sladění (*alignment*) lidských hodnot

a úmyslů s výkony AI a také diplomatického sladění lidí mezi sebou.

Henry Kissinger byl dvojicí svých spolupracovníků mentorem u druhé otázky, zatímco oni ho – díky zkušenostem získaným ve vedení Microsoftu a Googlu – prováděli první oblastí. Craig Mundie reprezentoval Microsoft na poli technologie při jednáních s americkou vládou i vládami dalších zemí a současně dohlížel na výzkumnou divizi firmy, načež se stal poradcem vedení výzkumné organizace OpenAI. Eric Schmidt vedl deset let společnost Google a dalších deset let věnoval vytváření sítě talentů a idejí na průsečíku technologie, filantropie a bezpečnosti s cílem chránit lidstvo před novými riziky.

Úkoly, jež před sebou spatřujeme, jsou vysoce naléhavé. Považujeme proto za nezbytné nečekat až na krizi, nýbrž zasazovat se už teď o to, aby se s nimi naše společnost i náš biologický druh vyrovnával proaktivně. A ačkoli je lidské bezpečí jednou nutnou složkou zdařilé odpovědi na AI, nedokáže zodpovědět všechny otázky, které s sebou AI nese, neboť lidstvo se v éře umělé inteligence promění. Otázkou zůstává, zda a do jaké míry se my lidé rozhodneme uchovat si moc nad tím, jak tato změna proběhne.

Část první

Na počátku

Kapitola 1

Objevitelství

Touha objevovat představuje možná vůbec nejopojnější vlastnost druhu *Homo sapiens*. Hnání zvědavostí a požitekem z překvapení vyplňujeme prázdná místa a proměňujeme kladené otázky v odpovědi. Průzkumné výboje jsou nedílnou součástí toho, jak sami sebe chápeme, a tak se vzdor zjevným hrozbám i zklamáním neúnavně ubíráme po četných cestách do neznáma.

Lidská lačnost zkoumat okolní fyzické prostředí vypráví především příběh ohromné odvahy tváří v tvář zdrcujícímu nebezpečí. Účastníci takovýchto výbojů naráželi na krajní limity existenčních možností. Když Fernão de Magalhães, řečený Magellan, počátkem 16. století obeplul zeměkouli, trvala jeho odysea tři roky a provázely

ji násilnosti, hladovění a smrt. Magellanova plavba jakožto první úspěšně zdolala trasu podél celé planety. Mořeplavec při ní suverénně překonal rekord ve výdrži na moři, stanovil rozměry Země, a jelikož se jeho výprava odehrála v kontextu evropského kolonialismu, prorazil cestu pro společenské a hospodářské kontakty v mezinárodním měřítku.

Většina Magellanových námořníků dobře věděla, že se vystavuje extrémnímu riziku. Obecně se sice už nevěřilo, že Země je plochá, avšak pro její kulatost zatím nebyl podán důkaz a mnozí členové posádky se nejspíš obávali, že přepadnou přes okraj světa.

Kapitán i posádka si uvědomovali, že pokud se ve svých předpokladech mylí, těžce za to zaplatí. Omylů se také skutečně dopustili a stihl je za to trest: neodhadli dobře délku plavby, množství proviantu ani hrozbu hromadných otrav, podcenili riziko toho, že se koráby poškodí nebo jinak vyřadí z provozu. Z původní flotily pěti lodí a zhruba 270 členů posádky se do španělského přístavu vrátil jen jeden jediný koráb s osmnácti vyhublými přízraky. Kapitán mezi nimi nebyl: zahynul cestou, když ho do nohy zasáhl otrávený šíp.¹

Během oněch čtyř set let, jež nás dělí od Magellanovy výpravy, byl svět zmapován do posledního místečka. Výjimku dlouho představovala Antarktida, světadíl stejně pustý jako vzdálené planety. K jižnímu pólu se více než kdokoli před ním přiblížil anglo-irský průzkumník Ernest Shackleton. S týmem bez sebemenších zkušeností a bez jakékoli státní podpory, pouze díky několika soukromým půjčkám a příspěvkům, dosáhlo jeho mužstvo

rekordu v délce takového putování a prozkoumalo terén po pozdější odvážlivce.

Shackleton sice jižního pólu nedosáhl, nicméně respekt pozdějších generací si vydobyl. Zajistil si ho tím, že namísto cestovatelských ambicí přiklkl prioritu lidským hodnotám. Rok po zahájení expedice členové jeho výpravy – den co den zapřažení do saní v brutálních desetidenních směnách – postupovali sotva o několik kilometrů denně. Vybavili se sice dostatkem jídla na to, aby se dostali až k pólu, jenže na cestu zpět k lodi jim už zbylé zásoby nepostačovaly.

Ve chvíli, kdy ho od triumfu dělilo sto padesát šest kilometrů, se Shackleton rozhodl neriskovat životy svých mužů a zavelel k návratu. „Učinili jsme vše, co bylo v našich silách,“ zaznamenal si do deníku.² Na zpáteční cestě svůj jediný denní suchar přenechával nemocnému členu výpravy Franku Wildovi a ten si pak do deníku zapsal: „Ten suchar by se nedal splatit žádnými penězi, které kdy vyšly z mincoven, a vzpomínka na tu obět je pro mne nesmazatelná.“³

Utrpěným nezdarem se Shackleton nenechal zastrašit a podnikl pak k jižnímu pólu ještě další expedice. Po desítky let se vyprávěla (následně vyvrácená) historika, že dal v londýnských *Timesech* zveřejnit inzerát následujícího znění:

Hledá se mužstvo na nebezpečnou výpravu. Nízká mzda, mráz, měsíce naprosté tmy, neustálé nebezpečí, návrat nejistý. V případě úspěchu pocty a uznání.⁴

Historika o tomto inzerátu je apokryfní, avšak obětavost, o níž vypovídá, přibarvená není. Přesně takto ještě před sto lety vypadala realita objevitelských cest: posouvání hranice možností záviselo na tom, kolik odvážlivců přistoupí na vysoce hazardní sázku.

Snad právě na důkaz uznání této odvahy se vlády některých států usnesly, že objevitelské výpravy podpoří a odmění, a tyto expedice se staly součástí mezinárodního soutěžení. Například Magellanovu plavbu definovala politika. Jelikož od vlastního panovníka – portugalského krále – finanční podporu neobdržel, prchl Magellan z vlasti a na výpravu vyplul díky ochotnější podpoře španělské koruny. Po jeho smrti si posádka za velitele zvolila Španěla Juana Sebastiána Elcana. Na zpáteční cestě se zoufalý Elcano takřka bez proviantu pokusil přistát u Kapverdských ostrovů, portugalské kolonie u západoafrického pobřeží, a vyslal na břeh třináct námořníků, aby záležitost dojednali s guvernérem. Došlo se jim však ponížujícího odmítnutí.

Elcano – zdrcený, ale o to odhodlanější předvést pošetilost Portugalska a úspěch Španělska – přikázal zvednout kotvu a pokračoval v plavbě do vlasti. (Vyjednávací výbor ponechal na nepřátelském území.) Když plavbu zdárně ukončil a splnil Magellanův sen, napsal španělskému králi, v té době již římskému císaři a nejmocnějšímu křesťanskému panovníkovi:

Vaše Výsost lépe než kdo jiný nahlédne, že tím, čeho si musíme vážit a cenit nejvíce, je skutečnost, že jsme objevili a obepluli celou oblohu světa a po vyplutí na západ jsme se navrátili od východu.⁵

Vážení čtenáři, právě jste dočetli ukázkou z knihy **Genesis** .

Pokud se Vám ukázka líbila, na našem webu si můžete zakoupit celou knihu.