

Uvnitř temné globální bitvy o zkrocení nejnebezpečnější technologie světa

 politico.eu/article/ai-control-kamala-harris-nick-clegg-meta-big-tech-social-media/

Může někdo ovládat AI?

Autor: **Mark Scott** ,
Gian Volpicelli ,
Mohar Chatterjee ,
Vincent Manancourt ,
Clothilde Goujard
a **Brendan Bordelon**

Foto-ilustrace Raphaëla Vicenziho pro POLITICO

V mokrém listopadovém odpoledni se americká viceprezidentka Kamala Harrisová a Nick Clegg z Mety doplazili do velkého stanu na pozemku anglického venkovského sídla z 19. století severně od Londýna, posadili se u kruhového stolu a mimo jiné se pustili do snaží zachránit lidstvo.

Uprostřed pošmourného počasí v Bletchley Park, který byl používán jako spojenecké velitelství pro prolamování kódů z 2. světové války, se Clegg a Harris připojili k elitnímu setkání globálních vůdců, akademiků a technologických manažerů, aby se vypořádali s tím, čeho se pesimisté obávají, je zničující nová hrozba, již čelí. svět: uprchlá umělá inteligence.

Politici a tech manažeři se po dvou dnech jednání dohodli na společném prohlášení o dobrých úmyslech, ale žádná jednotná odpověď na to, co by se mělo udělat. Namísto toho vytvořili soupeřící návrhy, jak řídit technologii, která bude dominovat většině příštího desetiletí – a pravděpodobně změní vše od obchodu a zdravotnictví až po samotnou demokracii.

Clegg, bývalý britský vicepremiér, tvrdil, že policejní AI se podobá stavbě letadla, které již letí – což je ze své podstaty riskantní a obtížná práce. Harris vytruboval snahy Washingtonu řešit nebezpečí AI prostřednictvím dobrovolných obchodních dohod jako světového zlatého standardu. Ursula von der Leyen, předsedkyně Evropské komise, která se také zúčastnila, vyzvala ostatní, aby se řídili novými, právně závaznými pravidly Bruselu a zakročili proti této technologii.

Debata představovala snímek větší pravdy. Poslední rok zuří po celém světě politický boj, většinou ve stínu, o to, jak – a zda – ovládat AI. Tato nová digitální skvělá hra ještě zdaleka nekončí. Kdokoli vyhraje, upevní svou nadvládu nad západními pravidly pro technologii definující éru. Jakmile budou tato pravidla nastavena, bude téměř nemožné je přepsat.

Pro ty, kdo sledovali konverzaci z první ruky, bylo smlouvání v britském dešti podobné evropským mocnostem v 19. století, které rozřezávaly svět.

„Připadalo mi to jako alternativní realita,“ řekla Amba Kak, ředitelka neziskové organizace AI Now Institute, která se diskusí účastnila. Na konci setkání podepsalo 29 zemí včetně Číny, členů Evropské unie a Spojených států dobrovolnou dohodu o snížení rizik, která se díky příchodu ChatGPT od OpenAI dostala do politického programu.

V nadcházejícím roce vytvoří boj o kontrolu technologie vítěze a poražené. Politici očekávají, že do konce roku 2024 bude dokončeno mnoho nových standardů umělé inteligence.

Pro tento článek POLITICO hovořilo s více než třemi desítkami politiků, politiků, technologických manažerů a dalších, z nichž mnozí získali anonymitu, aby mohli diskutovat o citlivých záležitostech, aby pochopili dynamiku ve hře, když se svět potýká s touto novou převratnou technologií.

Otázkou, před kterou stojí, je, zda USA, EU nebo Spojené království – nebo kdokoli jiný – bude schopen vymyslet plán, na kterém se západní demokracie dohodnou. Pokud se liberálním industrializovaným ekonomikám nepodaří dosáhnout společného režimu mezi sebou, může Čína zakročit, aby stanovila globální pravidla pro technologii, která – ve scénáři soudného dne – má určitý strach potenciál vymazat lidstvo z povrchu Země.

Zatímco Brusel a Washington předkládají své protichůdné plány na regulaci umělé inteligence, šance na dohodu nevypadají slibně.

Rytířství v Japonsku

Měsíc před konferencí v anglickém dešti se politici horečně snažili dosáhnout pokroku na druhém konci světa. Byl říjen a Věra Jourová vyčerpaná vystoupila ze svého 16hodinového letu z Bruselu do Japonska. Česká politička byla v nové roli nejvyššího zmocněnce EU pro technologie jen pár týdnů a její první mezinárodní angažmá nebude snadné.

Posláním Jourové bylo prodat evropský soubor pravidel umělé inteligence na schůzce G7, kde se sešli západní lídři, aby se pokusili navrhnout nové globální standardy pro nejpokročilejší formu této technologie, známou jako „generativní AI“, což je silný vývoj za ChatGPT a konkurenčními nástroji.



“Generative AI invaded our lives so quickly,” said Commissioner Věra Jourová. | Benoit Doppagne/Belga via AFP/Getty Images

Bruselský přístup je zakotven v zákonu EU o umělé inteligenci, prvním světovém pokusu o závaznou legislativu v této otázce. Na rozdíl od postoje, který upřednostňují USA, vize EU zahrnuje zákazy neinvazivnějších forem této technologie a přísná pravidla vyžadující, aby společnosti jako Google a Microsoft byly otevřenější ohledně toho, jak navrhují produkty založené na umělé inteligenci.

„Generativní umělá inteligence vtrhla do našich životů tak rychle a my potřebujeme něco rychlého,“ řekla Jourová POLITICO poté, co jela dvouhodinovým vlakem z Tokia do Kjóta na summit.

Na třídením setkání v Japonsku měl Nathaniel Fick hřiště soupeře.

Jako nejvyšší digitální diplomat Joe Bidena Fick, bývalý technický ředitel, nenavrhoval žádné zákazy ani přísné požadavky. Místo toho prosazoval lehčí režim založený většinou na dobrovolných závazcích průmyslu a stávajících domácích zákonech.

„Lidé mohou očekávat, že Spojené státy budou do všeho, co děláme, zaplétat otázky politiky AI,“ řekl Fick pro POLITICO po skončení summitu v Kjótu. "Rámce, kodexy, principy, které vyvíjíme, se stanou základem pro činnost."



Na duelech s tvůrci politik G7, vedoucími pracovníky technologických společností a dalšími vlivnými osobnostmi Jourová a Fick přednesli své případy.

Pro europoslankyni Jourovou byla situace jednoduchá. Brusel se již označil za digitálního policejního důstojníka Západu se záplavou předpisů na vše od ochrany soukromí spotřebitelů po zkrácení sociálních médií.

Časový rozvrh summitu byl nabitý, kromě několika přestávek na cigaretu a rychlých obědů v jídelně jen málo prostojů. Jourová tvrdila, že nezbytnou přísnost může zajistit pouze Evropa. EU by podle ní mohla firmám udělit pokuty za kasovní trháky a zakázat nejinvazivnější formy umělé inteligence – jako je sociální bodování, což jsou složité algoritmy sledující pohyb lidí, neslavně používané v Číně.

„Přišla s plánem, a tím nás měla přesvědčit, že evropská pravidla jsou jedinou hrou ve městě,“ řekl jeden z lidí, kteří se s Jourovou setkali. Další úředník ze země G7 řekl, že šéf digitálních technologií Evropy zamítl alternativní návrh Washingtonu pro nedostatek závazné legislativy.

Fickova protiofenzíva spoléhala na nespornou pozici Ameriky jako světové velmoci vývoje AI.

Politický pat na Capitol Hill znamená, že žádná komplexní legislativa z Washingtonu pravděpodobně v dohledné době nepřijde. Bílý dům už ale udělal spoustu oznámení. Loni v červenci Bidenova administrativa zajistila dobrovolné závazky předních technologických gigantů, jako je Amazon, aby byla AI bezpečnější. Poté Biden vydal exekutivní příkaz, oznámený 30. října, který zmocnil federální agentury jednat.



Americká viceprezidentka Kamala Harrisová zatroubila na úsilí Washingtonu řešit nebezpečí AI prostřednictvím dobrovolných obchodních dohod. | Drew Angerer/AFP přes Getty Images

Podle dvou osob, které se setkaly s Fickem v Japonsku, Fickovo vystoupení zahrnovalo kritiku bruselské legislativy za uvalení přílišné regulační zátěže na podniky ve srovnání s ochotou Washingtonu umožnit společnostem inovovat.

"Poselství bylo jasné," řekl další západní diplomat, který se zúčastnil summitu G7 v Kjótu. "Washington nedovolil Bruselu, aby se prosadil."

Dobré jídlo před apokalypsou

Zpátky v Evropě snahy o dosažení konsensu pokračovaly – v nějakém nepravděpodobném prostředí.

Okázalá večeře o šesti chodech v Elysejském paláci, oficiálním sídle francouzského prezidenta v Paříži z 18. století, zrovna nekřičí high-tech.

Ale během pětihodinového jídla gurmánské kuchyně a kvalitního vína v listopadu se Emmanuel Macron u soudu se 40 odborníky z oboru, včetně prezidenta OpenAI Grega Brockmana a hlavního guru AI společnosti Meta Yanna LeCuna.

Konverzace u večeře se rychle stočila k tvorbě pravidel.

Macron je globálním powerbrokerem. Francie má velké domácí odvětví umělé inteligence a prezident osobně touží vést mezinárodní úsilí o stanovení globálních pravidel pro tuto technologii. Na rozdíl od stereotypů francouzských politiků milujících regulaci zůstává Macron vůči bruselskému zákonu o AI opatrný, protože se obává, že to ochromí firmy jako Mistral, startup s umělou inteligencí, který spoluzaložil Cedric O, bývalý francouzský ministr pro digitální záležitosti.



Francouzský prezident Emmanuel Macron zůstává opatrný vůči bruselskému zákonu o AI. | Foto bazénu od Gonzala Fuentesese prostřednictvím AFP/Getty Images

Zatímco Macron pozorně poslouchal, hosté jeho večeře předložili své soupeřící návrhy. Diskuse odhalila další klíčový aspekt globální debaty o AI: střet mezi těmi, kteří věří, že nejzávažnější rizika jsou ještě mnoho let daleko, a těmi, kteří si myslí, že už tady jsou. Toto je bitva v jádru celé globální debaty. Staví proti sobě ty, kteří k potenciálu umělé inteligence přistupují optimističtěji a chtějí dovolit co nejvíce experimentů, proti ostatním, kteří se obávají, že už je příliš pozdě na to, aby chránili lidstvo před vážným poškozením.

Brockman z OpenAI, jeden z těch, kteří jsou uvolnění ohledně bezprostředních rizik a myslí si, že by se mělo zaměřit na řešení dlouhodobých hrozeb, řekl francouzskému prezidentovi, že AI je podle tří lidí, kteří se večeře zúčastnili, v drtivé většině silou dobra.

Jakákoli regulace – zejména pravidla, která by mohla brzdit raketový růst společnosti – by se měla zaměřit na dlouhodobé hrozby, jako je umělá inteligence, která nakonec převáží lidskou kontrolu, dodal.

Jiní jako Meredith Whittaker, americká technologička, která se také zúčastnila pařížské večeře v listopadu, tvrdili, že skutečné škody způsobené existující umělou inteligencí – včetně chybných souborů dat, které zakořeňují rasové a genderové předsudky – vyžadují, aby politici začali jednat hned.

"Macron to všechno vzal," řekl jeden z účastníků, který stejně jako ostatní přítomní dostal anonymitu, aby mohl o soukromé schůzce diskutovat. "Chtěl, aby lidé věděli, že Francie stojí za větší regulací, ale také, že Francie je otevřená i obchodu."

Podobné argumenty – staví proti potřebě pravidel, která by se zaměřovala na dlouhodobá rizika proti okamžitému nouzovému zásahu – se nyní objevují napříč západními hlavními městy.

Děsivé deepfakes

Zatímco se národní lídři snažili tento problém zabalit do hlavy, tech manažeři křížovali planetu a naléhali na politiky, aby neregulovali příliš.

Eric Schmidt, bývalý generální ředitel společnosti Google, a zakladatel LinkedIn Reid Hoffman letěli mezi Washingtonem, Londýnem a Paříží, aby nabídli svůj vlastní pohled na to, jak zacházet s AI, podle pěti osob obeznámených s diskusemi. Příliš horlivý zásah, tvrdili tech titáni, by poskytl autoritářské Číně rozhodující výhodu v AI. Mnozí jako Schmidt a Hoffman hodně investovali do startupů s umělou inteligencí.

Aby posílili své argumenty, techničtí šéfové, kteří chtějí regulaci s lehčím dotykem, přijali osobní. Lobbisté AI ukázali nejméně dvěma západním vůdcům realistická hluboce falešná videa – samotných vůdců. Tyto padělky generované umělou inteligencí jsou stále na špičce technologie a jsou často neohrabané a snadno odhalitelné. Podle tří osob, které byly na těchto schůzkách informovány, však lobbisté použili deepfakes, aby ukázali vůdcům, jak se tato technologie může vyvíjet, aby v budoucnu představovala vážná rizika.

Odpůrci tvrdí, že taková taktika umožňuje společnostem jako OpenAI a Alphabet podceňovat, jak tato technologie právě teď poškozuje lidi – například tím, že odmítají žádosti o státní výhody od menšin a znevýhodněných lidí ve společnosti.

"Většina Američanů si neuvědomuje, že umělá inteligence už existuje," řekl Cory Booker, demokratický americký senátor. "Chci se ujistit, že naše současné zákony proti diskriminaci... naše současné zákony potvrzující ochranu jsou skutečně vymáhány."

V roce 2024 se zdá, že vyhrávají ti, kteří chtějí lehčí dotek, a to navzdory novým závazným pravidlům EU pro umělou inteligenci.



Eric Schmidt, bývalý generální ředitel společnosti Google, zakladatel LinkedIn Reid Hoffman a Sam Altman z OpenAI byli mezi těmi, kteří se v Londýně připojili ke světovým lídrům, aby v loňském roce diskutovali o AI. | Leon Neal/Getty Images

Poté, co Brusel v prosinci dosáhl politické dohody o svém zákonu o AI, Macron – čerstvě po své osobní večeři o AI – se postavil na stranu německého kancléře Olafa Scholze a italského premiéra Giorgia Meloniho, aby požadovali méně kontrol nejnovějších forem technologie. Francouzský prezident se obával, že taková regulace bude překážet místním šampionům – včetně Mistralu, francouzské společnosti pro umělou inteligenci, kterou částečně podporuje Schmidt, bývalý šéf Googlu.

Francie zůstala zdrženlivá, až nakonec na začátku února podlehla politickému tlaku, aby pravidla podpořila, i když s velkými výhradami. „Budeme regulovat věci, které už nebudeme vyrábět nebo vymýšlet,“ řekl Macron publiku v Toulouse poté, co zajistil evropským firmám na poslední chvíli vyškrtnutí. "Tohle není nikdy dobrý nápad."

Biozbraně a velká technika

Tristan Harris si myslel, že zasáhl knockoutem.

Na soukromém slyšení mezi americkými zákonodárci a technologickými experty v září Harris, spoluzakladatel neziskové organizace Center for Humane Technology, popsal, jak jeho inženýři donutili nejnovější produkt AI společnosti Meta, aby spáchal děsivý čin: konstrukci biologické zbraně. .

To ukázalo, že přístup Meta k umělé inteligenci byl příliš laxní a potenciálně nebezpečný, tvrdil Harris, podle dvou úředníků, kteří se setkání zúčastnili.

Technologický gigant Marka Zuckerberga upřednostňoval takzvanou open-source technologii – umělou inteligenci snadno dostupnou pro všechny – s několika zárukami proti zneužití. Taková otevřenost, dodal Harris, by vedla ke škodám v reálném světě, včetně šíření zbraní hromadného ničení generovaných umělou inteligencí.



Na rozdíl od jiných gigantů ze Silicon Valley se Meta Marka Zuckerberga postavila na stranu těch, kteří chtějí technologii ponechat otevřenou pro všechny. | Josh Edelson/AFP přes Getty Images

Jeho triumf na sebe nenechal dlouho čekat. Zuckerberg, který byl také přítomen slyšení v Capitol Hill, rychle vytáhl svůj telefon a našel stejné informace o biologických zbraních pomocí jednoduchého vyhledávání Google. Zuckerbergův úder vedl k smíchu z místnosti. Oslabilo to Harrisovo obvinění, že přístup Meta s otevřeným zdrojovým kódem AI byl hrozbou pro lidstvo.

"Nebyl to ten slam dunk, jak si Harris myslel," řekl jeden z úředníků, kterým byla udělena anonymita, aby mohli o schůzce diskutovat.

Zúčtování Harrise-Zuckerberga zachycuje další klíčovou otázku pro tvůrce politik: Měla by být umělá inteligence omezena na hrstku důvěryhodných společností nebo by se měla otevřít všem příchozím k využití?

Podle těch, kteří upřednostňují licencování pokročilé umělé inteligence několika vybraným společnostem, jsou bezprostřední rizika příliš vážná na to, aby se pustili do širšího světa. Prezentace v PowerPointu varující, že modely umělé inteligence s otevřeným zdrojovým kódem by uvolnily nekontrolovatelné biologické zbraně přístupné

teroristům, byla soukromě prezentována vládním úředníkům v Londýně, Paříži a Washingtonu, podle šesti politiků zapojených do těchto setkání. POLITICO nedokázalo určit, která organizace stojí za těmito hrozivými předpověďmi.

Microsoft a OpenAI patří mezi společnosti, které upřednostňují omezení této technologie na malý počet firem, aby si regulační orgány mohly vybudovat vazby s inovátory AI.

„Přístup založený na licencích je správnou cestou vpřed,“ řekla společnosti POLITICO Natasha Crampton, hlavní odpovědná AI společnosti Microsoft. "Umožňuje úzkou interakci mezi vývojářem technologie a regulátorem, aby bylo možné skutečně posoudit rizika."



Tristan Harris, spoluzakladatel Centra pro humánní technologie, vystoupí v New Yorku v roce 2022. | Craig Barritt/Getty Images pro Unfinished Live

Kritici dominance Silicon Valley tvrdí, že Big Tech má zájem na vyloučení konkurence. Podle obhájce open source Marka Surmana, výkonného ředitele Mozilla Foundation, chtějí největší světové technologické společnosti zachytit začínající průmysl umělé inteligence pro sebe a odříznout konkurenční nováčky.

Během podzimu Surman cestoval po západních hlavních městech, aby vyzval vysoké představitele USA, EU a Spojeného království, aby neomezovali, kdo může vyvíjet modely umělé inteligence nové generace. Na rozdíl od jiných gigantů ze Silicon Valley se Zuckerbergova Meta postavila na stranu lidí, jako je Surman, kteří chtějí mít technologii otevřenou pro všechny, což posílil svým střetnutím s Harrisem ohledně biologických zbraní v září.

Otevřený přístup si získává přátele v mocných částech Washingtonu, ačkoli američtí představitelé zůstávají rozděleni. V Bruselu se tvůrci politik dohodli, že téměř veškerou umělou inteligenci s otevřeným zdrojovým kódem osvobodí od nejpřísnějšího dohledu – pokud se tato technologie nepoužívá v nezákonných praktikách, jako je velkoobchodní rozpoznávání obličejů. Ve Spojeném království se však úředníci přiklánějí k omezení nejpokročilejší umělé inteligence na hrstku zemí.

"Pokud necháme AI v rukou několika málo lidí, nebude tam dost očí, které by zkoumaly, co s ní dělají," řekl Surman.

Budoucnost je tady

Pokud byl rok 2023 rokem, kdy lobování AI proniklo do hlavního politického proudu, rok 2024 rozhodne, kdo vyhraje.

Klíčová data rozhodnutí se rychle blíží. Do léta vstoupí v platnost částí evropského zákona o umělé inteligenci. Konkurenční výkonný příkaz AI z Bílého domu již mění způsob, jakým federální agentury zacházejí s technologií. Další reformy ve Washingtonu se očekávají do konce roku.

Další země jako Spojené království, Japonsko a Kanada – stejně jako Čína – připravují své vlastní plány, které zahrnují větší dohled nad nejpokročilejší AI a snahy přimět ostatní země, aby přijaly jejich pravidla.

Podle desítek odborníků na umělou inteligenci, kteří hovořili s POLITICO, je jasné, že v roce 2024 touha veřejnosti po politické akci pravděpodobně nezmizí. Dopad umělé inteligence, zejména v roce masových voleb z USA a EU přes Filipíny až po Indii, není jistý.

„Tato záležitost o ztrátě kontroly nad umělou inteligencí je náhle brána vážně,“ řekl Max Tegmark, prezident Future of Life Institute, neziskové organizace, která obhájí větší kontrolu AI. „Lidé si uvědomují, že to už není dlouhodobá záležitost. Děje se to.“

Mark Scott hlásil z Londýna, Bruselu, Paříže, Berlína a Washingtonu; Clothilde Goujard z Kjóta, Japonsko, Vincent Manancourt z Londýna, Gian Volpicelli z Bruselu, Mohar Chatterjee a Brendan Bordelon z Washingtonu .

OPRAVA: Předchozí verze tohoto článku chybně uvedla, které firmy upřednostňují omezení technologie na malý počet firem. Jsou to Microsoft a OpenAI.