FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

KATEDRA
KYBERNETIKY

Dario Amodei



"There's a 25% chance that things go really, really badly." -
Dario Amodei na p(doom) otázku.

Západočeská Univerzita V Plzni
Katedra Kybernetiky
Semestrální práce - HKUI

David Straka
1. semestr
25. listopadu 2025

1 Životopis

1.1 Raný život a vzdělání

Dario Amodei se narodil v roce 1983 v San Franciscu (Kalifornie, USA). Jeho otcem byl Riccardo Amodei - italsko-americký řemeslník, matkou Elena Engel - americká židovka, která byla projektová manažerka knihovních staveb a renovací. V roce 2006 zemřel jeho otec po dlouhé nemoci, což podle Daria výrazně ovlivnilo směr jeho budoucího akademického a výzkumného zájmu.

Od dětství byl Dario silně zaměřen na matematiku a fyziku: v jednom ze svých rozhovorů připomíná, že ho spíše zajímalo objevování fundamentálních vědeckých pravd než budování webových stránek.

Studoval střední školu v San Franciscu a následně začal studovat na Stanford University, kde získal bakalářský titul. Poté pokračoval na doktorandské studium na Princeton University, kde se věnoval biofyzice a fyzice. Po získání PhD absolvoval post-doktorandské působení na Stanford University School of Medicine, kde se zabýval aplikacemi hmotnostní spektrometrie na síťové modely buněčného proteomu a hledáním biomarkerů rakoviny.

Dario byl také oceněn Hertzovou nadací, což je prestižní stipendijní program pro mladé výzkumníky.

Tato kombinace studia - od fyziky, přes neurovědu až po biofyziku - vytvořila základ jeho interdisciplinárního myšlení a později napomohla i přechodu k výzkumu v umělé inteligenci.

1.2 Výzkumná dráha

Během post-doktorského studia narazil na lidské výzkumné limity a viděl velký potenciál v rozvíjejících se AI technologiích. Po ukončení studia přešel z akademického prostředí do výzkumného světa AI a softwaru v korporátním světě, jelikož vnímal velké investice od těchto komerčních firem.

Zpočátku jeho kariéry přemýšlel nad vytvořením startupu. Od tohoto záměru ale upustil. Již v roce 2014 začal pracovat v technologickém sektoru: například u firmy Baidu (Čína) nebo také ve firmě Google v jejich výzkumném týmu Google Brain. V této době již začínal pozorovat, jak nárůst výpočetní kapacity a dat vede k prudkému zlepšování modelů - tzv. "scaling laws", které se později staly klíčovým konceptem v oblasti velkých jazykových modelů.

Tyto zkušenosti mu poskytly odbornou i strategickou perspektivu. Pochopil, že umělá inteligence není pouze o algoritmech, ale také o infrastruktuře, měřítku, spolupráci a bezpečnosti.

1.3 OpenAI a cesta k podnikání

V roce 2016 se Dario Amodei připojil k OpenAI, kde postupně zastával významné role v oddělení výzkumu - nakonec jako vice-prezident výzkumu (Vice President of Research). V této pozici se podílel na vývoji průlomových modelů GPT-2 a GPT-3 (velké jazykové modely) a též na výzkumu bezpečnosti umělé inteligence.

V průběhu svého působení v OpenAI začal mít rozdílné názory na strategii vývoje AI: podle jeho slov si přál větší zaměření na bezpečnost, řízení rizik, a kontrolu měřítka, než na pouhý výkon a rychlé uvedení na trh. Tyto rozdíly nakonec vyústily v odchod několika členů týmu a v založení nové společnosti.

V roce 2021 spolu s jeho sestrou Danielou Amodei a skupinou bývalých kolegů opustil OpenAI a založil společnost Anthropic.

1.4 Založení a vedení Anthropic

Vznik Anthropic v roce 2021 byl do značné míry odrazem intelektuální vize a profesní dráhy Daria Amodeie. Ten po odchodu z OpenAI založil společnost s jasně formulovaným cílem: vyvíjet systémy umělé inteligence, které budou nejen technologicky výkonné, ale zároveň bezpečné. Amodei jako CEO vtiskl společnosti charakter výzkumné instituce orientované na dlouhodobý veřejný prospěch. Opakovaně zdůrazňuje, že vývoj AI by neměl být veden pouze komerčními ambicemi, ale měl by reflektovat rizika a dopady na společnost, ekonomiku i demokracii.

Pod jeho vedením si firma Anthropic rychle získala mezinárodní viditelnost a začala uvádět modely Claude, které se staly jedním z hlavních referenčních bodů pro diskuzi o bezpečných velkých modelech. Ať už jde o první generaci Claude 1 a Instant, výrazně posílené Claude 2.0 a 2.1 s širokým kontextem, nebo novou řadu Claude 3 a Claude 4 (s verzemi sonnet, opus a haiku), všechny tyto modely reflektují Amodeiův důraz na kombinaci vysokého výkonu a robustních bezpečnostních protokolů.

1.5 Myšlenky, vliv a současné výzvy

Dario Amodei patří k nejvýznamnějším osobnostem debaty o budoucnosti umělé inteligence. Jeho veřejná stanoviska i výzkumné aktivity ukazují několik zásadních okruhů, které dlouhodobě formují diskusi v akademické i technologické komunitě:

- Riziko škálování AI: Amodei upozorňuje, že exponenciální růst výpočetních kapacit a dat může vést ke vzniku systémů s vlastnostmi, které překračují lidskou intuici i kontrolu. V rozhovorech uvádí, že nevidí žádné známky zpomalování technického pokroku. Právě proto považuje bezpečnostní výzkum za otázku naléhavé priority, nikoli za doprovodnou aktivitu.
- Bezpečnost a ovladatelnost modelů: Základní premisou jeho přístupu je, že umělá inteligence musí zůstat řiditelná i v extrémních scénářích. Amodei zdůrazňuje, že systémy je nutné navrhovat tak, aby bylo možné spolehlivě řídit jejich chování a aby modely konzistentně sledovaly instrukce v souladu s očekávanými cíli.
- Ekonomické a společenské dopady: Amodei pravidelně upozorňuje na to, že AI není pouze technologický nástroj, ale systémový faktor s výrazným dopadem na pracovní trh, ekonomickou stabilitu i mediální prostředí. Změny, které může přinést vyžadují připravenost vlád i firem.
- Úloha demokracií vývoji AI: Amodei patří k zastáncům názoru, že demokratické státy musí hrát aktivní roli ve výzkumu a regulaci pokročilé AI. Nejde podle něj pouze o otázku bezpečnosti, ale i o geopolitickou stabilitu a zabránění monopolizaci moci technologickými giganty.

1.6 Závěr

Dario Amodei je jednou z klíčových osobností současného odvětví umělé inteligence nejen díky svému vedení firmy Anthropic, ale také díky důrazu na bezpečnost a dlouhodobé řízení rizik.

Firma Anthropic je pod jeho vedením schopna soutěžit s největšími hráči v oboru a zároveň skrze ni uvádí v život svou vizi, že vývoj AI musí být stejně tak otázkou etiky a odpovědnosti jako otázkou inovací.

Pokud bude budoucnost AI skutečně určována schopností propojit sílu modelů s jejich bezpečným a kontrolovaným chováním, pak je Amodeiho přístup jedním z klíčových směrů, které tuto budoucnost už nyní vytvářejí.

Reference

[1] Dario Amodei. *Personal website.* Available at: https://www.darioamodei.com

[2] Time Magazine. *Anthropic CEO Dario Amodei on the Future of AI.* Available at: https://time.com/6990386/anthropic-dario-amodei-interview

[3] Hertz Foundation. *Dario Amodei.* Available at: https://www.hertzfoundation.org/people/dario-amodei

[4] Kantrowitz, A. *The Making of Anthropic CEO Dario Amodei.* Medium, 2023. Available at: https://kantrowitz.medium.com/the-making-of-anthropic-ceo-dario-amodei-449777529dd6

[5] Financial Times. *Dario Amodei: The AI Safety Pioneer Shaping the Industry's Future.* Available at: https://www.ft.com/content/e75e3388-4700-413d-ab67-778410c2d977

[6] Axios. *Anthropic CEO Dario Amodei says there is a 25% chance of "P(doom)".* Available at: https://www.axios.com/2025/09/17/anthropic-dario-amodei-p-doom-25-percent

[7] Anthropic. *Claude Model and API Documentation.* Available at: https://platform.claude.com/docs