

# Stanovisko CERAI k vybraným otázkam MIRRI k návrhu AIA

Bratislava, 26.4.2022

Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky (MIRRI) požiadalo Stálu komisiu pre etiku a reguláciu umelej inteligencie (CERAI) o vyjadrenie k AIA. Na 6. riadnom zasadnutí CERAI konanom dňa 26.4.2022 členovia CERAI diskutovali dve otázky z oblasti neuromorfných systémov a neurotechnológií. Na základe materiálu predloženého vybranými členmi CERAI a následnej diskusie pripravila Stála komisia spoločné stanovisko k položeným otázkam. Toto stanovisko bolo následne schválené per rollam. Na základe výsledkov hlasovania sa Stála komisia rozhodla vydať nasledujúce stanovisko.

## **Otázka 1: Pokrýva návrh AIA dostatočne, resp. vôbec, tzv. neuromorfné systémy, a je vhodné ich už zahrnúť do regulácie?**

Stanovisko:

Komisia má za to, že v aktuálnom návrhu AIA je otázka tzv. neuromorfných systémov nepokrytá, pričom už aj akademická obec upozorňuje na nekonceptný prístup vlád a regulátorov v tejto oblasti<sup>1</sup> Navrhujeme preto túto oblasť naďalej monitorovať, resp. podporiť globálne trendy v oblasti regulácie.

Návrh AIA predkladá reguláciu pre tradičné systémy s vopred daným a jasne definovaným datasetom a režimom tréovania. Tradičný AI system využíva dataset a log, kde sa štatistickými a analytickými metódami vypočítava model neurónovej siete. Neuromorfné systémy svoju neurónovú sieť formujú za pomoci spätnej väzby samotnej neurónovej siete, pričom negenerujú tradičný log. Tradičný AI systém je teda z tohto pohľadu štruktúralne statický, pričom štruktúra jeho neurónovej siete je definovaná deterministickým výpočtom z datasetu. Neuromorfný systém je štruktúralne plastický, pričom štruktúra jeho neurónovej siete je definovaná vstupom ovplyvneným hlavne prostredím, kde je nasadený. Čiže dva identické neuromorfné systémy nasadené v rôznych prostrediach budú mať rôznu konfiguráciu neurónovej siete, keďže prostredie a spätná väzba tu nahrádzajú potrebu datasetu. Keďže AIA sa snaží ošetriť reguláciu najmä cez logy a dataset, toto bude pri neuromorfných systémoch predstavovať závažný problém.

Neuromorfné systémy sa líšia od tradičných "deep learning AI" systémov hlavne v oblasti domény použitia. Tam kde je vyžadovaná presnosť kalkulácie, spoľahlivosť a deterministickosť výpočtu, tam sa budú zrejme používať digitálne výpočty (realizované klasickými neurónovými sieťami). Neuromorfné systémy počítajú analógovo a preto sú oveľa energeticky efektívnejšie. Sú vhodné najmä na priamu interakciu v neštruktúrovanom a zašumenom prostredí zaťaženom neurčitou.

---

<sup>1</sup> <https://www.nature.com/articles/s41586-021-04362-w>

## Otázka 2: Poskytuje AIA dostatočnú, resp. vôbec nejakú, ochranu pred neurotechnologiami?

Stanovisko:

AIA neposkytuje prakticky žiadnu (resp. minimálnu) ochranu pre neurotechnologiami. Report vypracovaný pre DH-BIO<sup>2</sup> považujeme prakticky za nezpracovaný a pohliadajúci na tému z biomedicinskeho hľadiska. Považujeme za dôležité zapracovať do návrhu AIA aj technický pohľad a mieru aktuálneho technologického pokroku v tejto oblasti.

Neurotechnologie možno vnímať ako AI zariadenia, ktoré sa prispôbujú prostrediu na základe interakcie s ním. Možno ich rozdeliť na systémy s nebiologickou interakciou (voľne EQXX mercedes koncepte, DB optimalizácia vlakov,....) a biologickou interakciou (mozgovo počítačový interface, umelá synapsia). Oboje svoju štruktúru upravujú na základe prostredia, resp. užívateľa, čo pri nebiologickej interakcii môže viesť ku zmene, resp. neplánovanému rozšíreniu funkcionality, ktorá môže v zariadení, kde sú implementované viesť k nepredpokladaným stavom, ktoré môžu ovplyvniť mechanickú funkčnosť či bezpečnosť zariadenia. Pri neuromorfných systémoch s priamou biologickou interakciou, je sporná otázka ovplyvňovania vedomia a užívateľa, keďže elektrická reakcia mozgu je impulzom pre stav vedomia človeka, čiže je možné ovplyvniť stav vedomia jedinca bez jeho vedomia.

Taktiež neuromorfné technológie zapríčiňujú zmeny v prirodzenej štruktúre vývoja neurónovej siete v ľudskom mozgu. Otázka je ochrana vedomia jedinca, lebo po odobratí mozgovo počítačového interface po dlhšej dobe omnoho synaptických spojení a axonov stáva slepými uličkami, čo vedie k závažným poškodeniam mozgu a vedomia jedinca. Je dôležité myslieť na štandardy, kvôli udržateľnosti BCI a aj na medicínske účely a iné účely kvôli zachovaniu mentálnej integrity jedinca.

Ďalšou významnou témou je biometrická identifikácia na báze snímania mozgových vln bezkontaktnou metódou, ide o najunikatejšiu biometriu doposiaľ objavenú, ale prichádza s rizikami ovplyvňovania vedomia identifikovaného jedinca. Samostatnou kapitolou sú biologicky kompatibilné mozgovo počítačové rozhrania, a metódy ich integrácie na báze traumatických zrazení mozgu (dlhých axónov), ktoré majú potenciál meniť vedomie, povahu a chovanie jedinca už po veľmi krátkom čase použitia rádovo dní (nezvratné zmeny), pričom biologické zmeny v štruktúre mozgu boli pozorované už po pár hodinách od ich integrácie.

Viac sa touto problematikou zaoberá napr. voľne dostupná kniha MacKellar, C. (2019): *Cyborg Mind: What Brain–Computer and Mind–Cyberspace Interfaces Mean for Cyberneuroethics*<sup>3</sup>, v ktorej autori analyzujú potenciálne problémy s neurotechnologiami (Changing Cognition and Consciousness; Escaping Reality; Changing Mood, Personality, Identity; Uploading a Mind; Issues of Privacy; New Cybercrimes; Human autonomy). V závere knihy (s 239-242) sú odporúčania pre cyberneuroetiku, ktoré vypracoval Scottish Council on Human Bioethics - akási Charta základných ľudských práv aktualizovaná o ohrozenia a príležitosti spojené s neurotechnologiami.

---

<sup>2</sup> <https://rm.coe.int/report-final-en/1680a429f3>

<sup>3</sup> <https://www.berghahnbooks.com/title/MacKellarCyborg>