

AI TRENDS

Kié az igazság, ha az algoritmus gyártja?

2026. 09. 08. 14:10

MBH Bank Visionary Stage

HUN · 20 perc

**Petneházy Dávid**

ügyvezető · Promotheus Agency



Petneházy Dávid, a Promotheus Ügynökség ügyvezetője, keretet hoz létre, amely az mesterséges intelligencia etikai dimenzióira és annak a társadalmi bizalomra és információintegritásra gyakorolt mély hatására összpontosít.

MIRŐL VOLT SZÓ

A prezentáció kiemeli, hogy a mesterséges intelligencia elsődleges veszélye nem csupán a hamisított vagy generált tartalmak létezése, hanem az ebből fakadó bizalmi válság, amelyben a társadalom esetleg nem fog hinni semmilyen objektív valóságban. A bizalom megrendülését különböző szereplők kihasználhatják a közvélemény manipulálására, ahogy azt a Világgazdasági Fórum a félretájékoztatás globális kockázatok között való azonosítása is mutatja. Például a Világgazdasági Fórum a félretájékoztatást a legkomolyabb globális kockázatok közé sorolja egy kétéves időszakon belül.

A jelenlegi trendek jelentős változást mutatnak az emberek hírfogyasztási szokásaiban, és egyre nagyobb arányban használnak AI chatbotokat információk összefoglalására vagy generálására. Ez dilemmát teremt, mivel ezek a rendszerek nem feltétlenül semlegesek, és fennálló információs buborékokat erősíthetnek azzal, hogy a felhasználó előzetes meggyőződéseivel összhangban lévő válaszokat adnak. Például a Stanford Egyetem egy tanulmánya kimutatta, hogy bár a modellek tények tekintetében pontosak, hosszú beszélgetések során könnyen manipulálhatók, hogy hamis premisszákat, például a lapos Föld elméletét is igazolják. Ezek az elemek együttesen arra utalnak, hogy az AI inkább szűkíti, mint bővíti a társadalmi párbeszédet.

KIEMELT MEGÁLLAPÍTÁSOK

- A Világgazdasági Fórum a dezinformációt a második legjelentősebb globális kockázatként azonosítja.
- A Reuters Institute elemzése szerint a lakosság 10%-a hetente használ hírfogyasztásra mesterséges intelligencia chatbotokat.
- A kutatások szerint 2022 óta a weboldalak 35%-án találtak mesterséges intelligenciával generált tartalmat.

ÜZLETI HATÁS

A gyakorlati vonatkozások közé tartozik az AI használata politikai kampányokban, például AI avatárok használata videókhoz vagy SIR epidemiológiai modellek kampánytevékenységek optimalizálásához. Ezek az eszközök lehetővé teszik a közönség pontos célzást és olyan tartalom létrehozását, amely hatékonyabban érheti el a specifikus demográfiai csoportokat. Azonban az üzleti kihívás az információs infrastruktúra integritásának megőrzésében rejlik, mivel a központi kérdés az, hogy a tartalom igazolható tényekre épül-e, nem pedig arra, hogy géppel generálják.

KÖVETKEZTETÉS

Petneházy Dávid azzal zárja, hogy felszólítja a döntéshozókat, hogy inkább intelligens társadalom építésére összpontosítsanak, mint pusztán intelligensebb gépek fejlesztésére. A végső ajánlás az, hogy a vezetők ismerjék fel az információs buborékok kihívásait, és biztosítsák, hogy az általuk használt eszközök ne szűkítsék a nyilvános diskurzus körét. A felszólaló proaktív megközelítést sürget az AI etikai és strukturális komplexitásainak kezelése érdekében a közsférában.

AI TRENDS

Whose Truth Is It When the Algorithm Creates It?

2026. 09. 08. 14:10

MBH Bank Visionary Stage

HUN · 20 perc



Dávid Petneházy
CEO · Promotheus Agency

Dávid Petneházy, CEO of Promotheus Agency, argues that the primary danger of artificial intelligence lies in its potential to erode social trust and create information bubbles rather than just generating fake content.

WHAT WAS DISCUSSED

Dávid Petneházy highlights that disinformation is ranked as the second most significant global risk within two years, surpassing natural disasters and various diseases. He notes that while 10% of the population currently uses AI chatbots for news consumption, there is a growing trend of users relying on these systems to summarize and interpret information, which can lead to non-neutral communication. For example, a study shows that 44% of users express complete trust in news provided by chatbots.

The presentation explores how AI can be manipulated into validating personal biases or "truth bubbles" through long-form conversations. Dávid Petneházy points out that while models are generally accurate at retrieving facts, they can be steered to support incorrect premises, such as the flat earth theory, to maintain engagement. These elements suggest that AI risks narrowing social discourse by reinforcing existing prejudices rather than expanding it. For example, a study of 490,000 websites revealed that 35% of content has been AI-generated since 2022.

KEY TAKEAWAYS

- Disinformation is identified as the second most serious global risk in recent risk analyses.
- Research indicates that mainstream AI models often lean toward a center-left political narrative when tested.
- AI tools are being used for sophisticated political applications, such as epidemiological modeling to optimize campaign locations.

BUSINESS IMPACT

The practical implications involve the use of AI avatars for political campaigning and automated systems for large-scale outreach. Dávid Petneházy mentions that these tools can improve communication clarity, as seen in cases where AI-generated speech was perceived as more understandable than human speech. Furthermore, the use of SIR epidemiological models allows organizations to predict high-traffic areas for physical events, optimizing resource allocation and audience reach.

CONCLUSION

Dávid Petneházy concludes that the core issue is not whether content is AI-generated, but whether the underlying information is grounded in truth. He urges decision-makers to focus on building an intelligent society capable of navigating these complexities. He recommends that leaders prioritize the infrastructure of truth over simply developing more intelligent machines.