

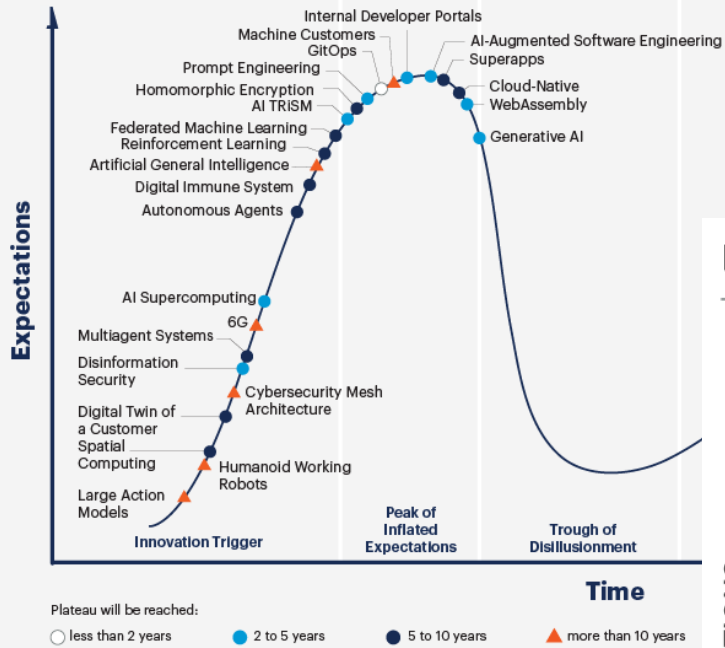
Hová megy ez a vonat? helyzetkép

conTEXT konferencia 2024

Körmendi György

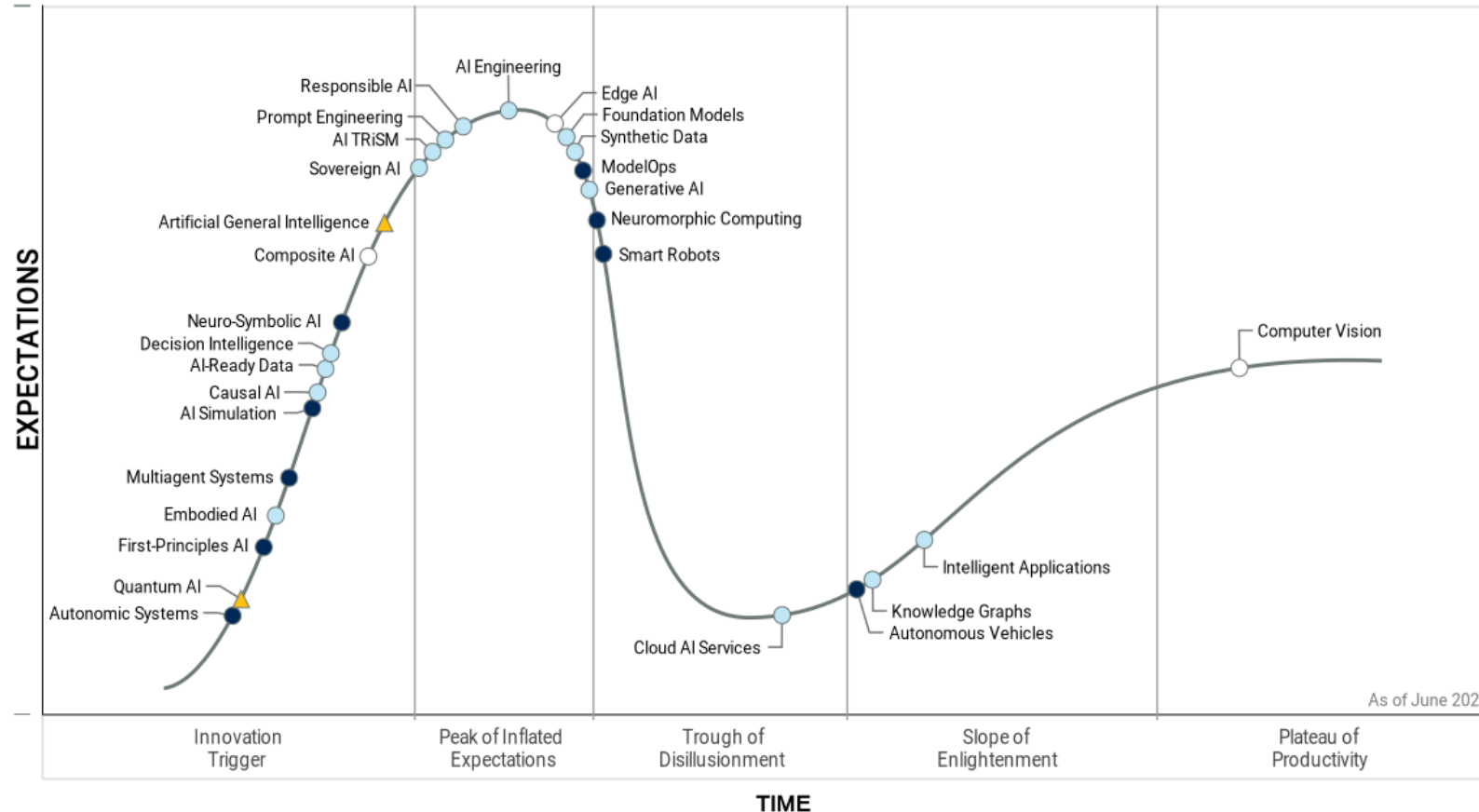


Hype Cycle for Emerging Technologies, 2024



Source: Gartner
 Commercial reuse requires approval from Gartner and must comply with the Gartner Content Compliance Policy
 © 2024 Gartner, Inc. and/or its affiliates. All rights reserved. 3205434

Hype Cycle for Artificial Intelligence, 2024

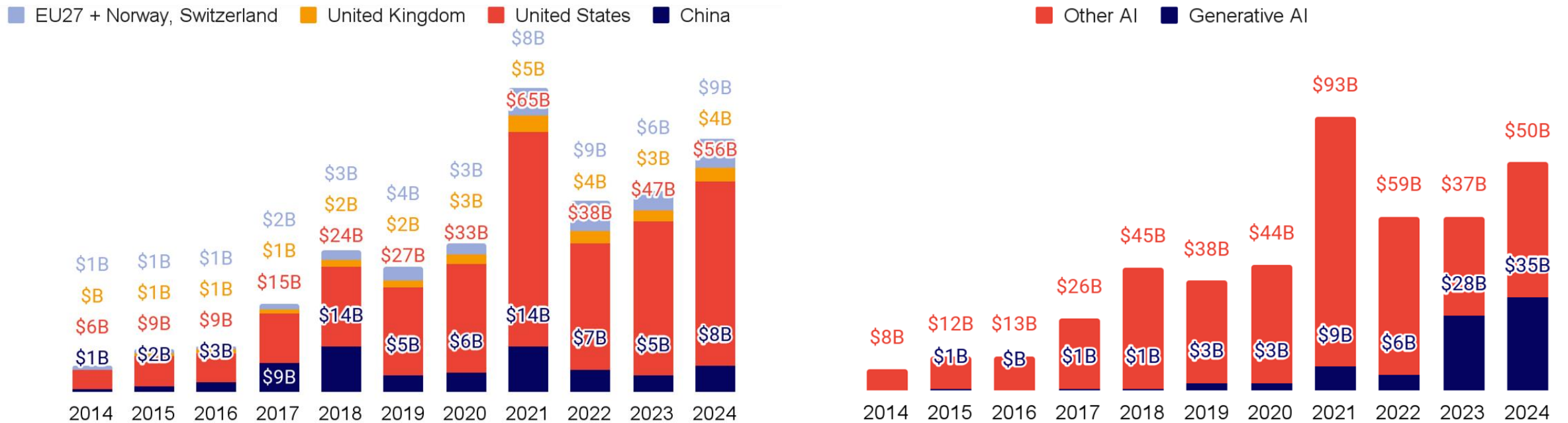


As of June 2024



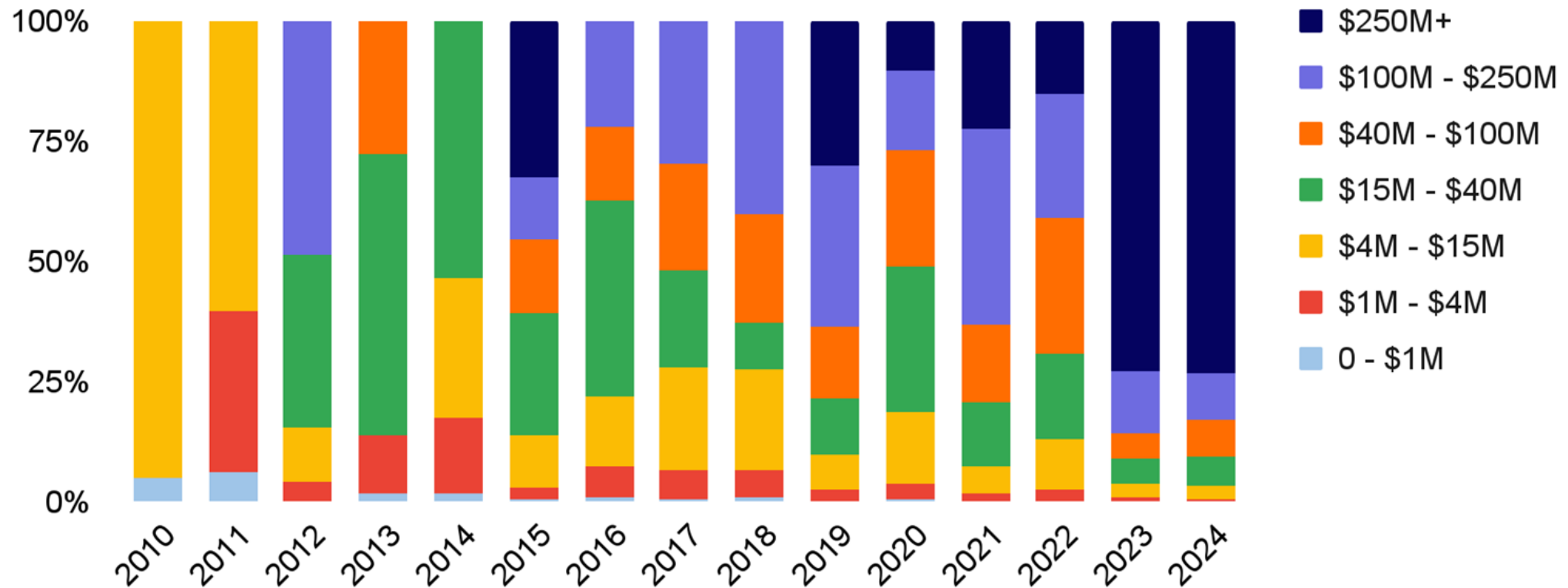
AI investment surges in every region

▶ Driven by GenAI mega-rounds like xAI and OpenAI's \$6B fundraises, US private market continue to lead. Total investment into AI companies reached close to \$100B.



While over the last 2 years, mega \$250M+ rounds dominated AI financings

- ▶ There appears to be a clear “pre/post-GPT-4 era” (2023) that triggered all funding systems to go on steroids...



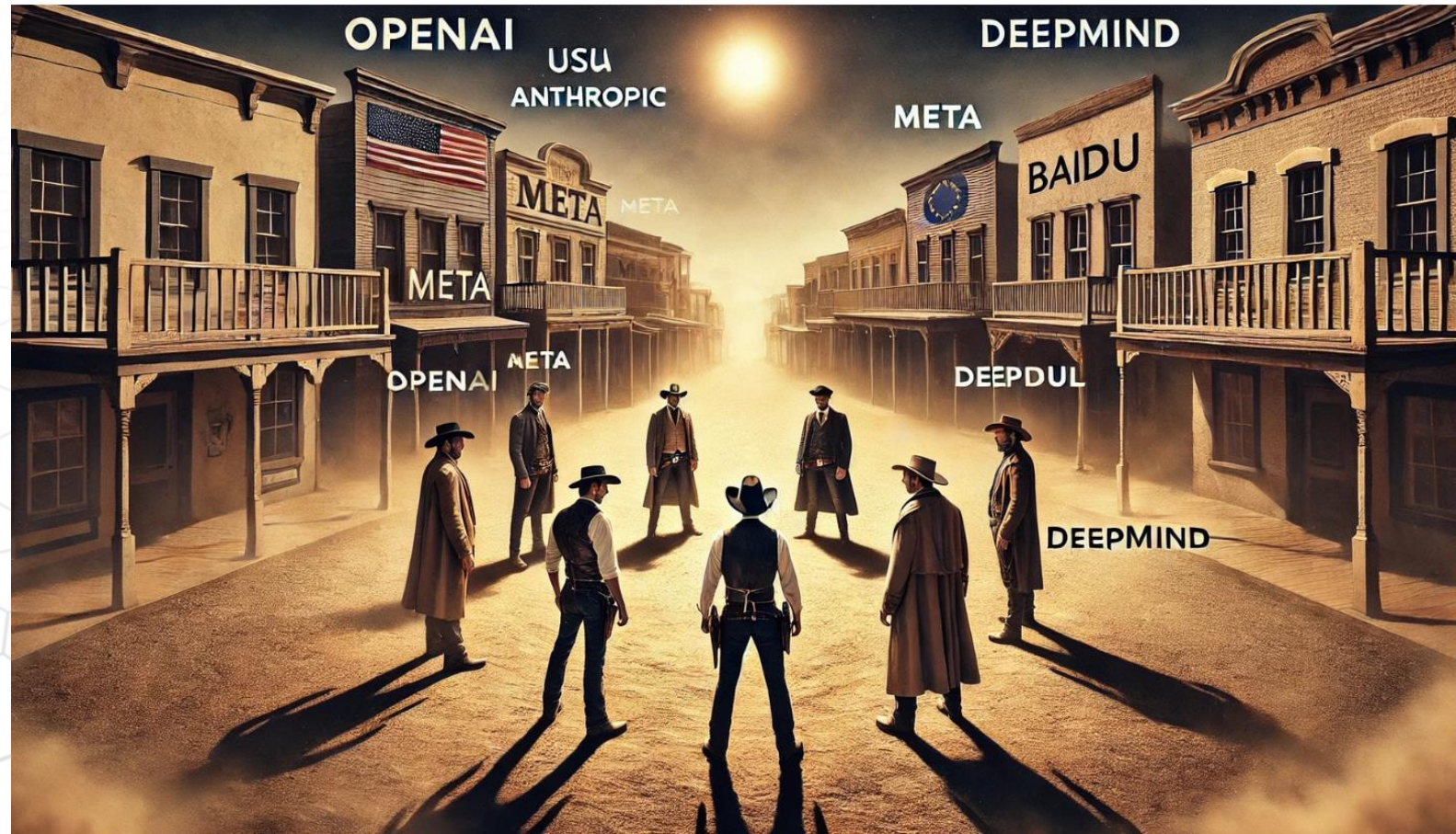
OPEN SOURCE... HMMM

Project	Availability						Documentation				Access			
	Open code	LLM data	LLM weights	RL data	RL weights	License	Code	Architecture	Preprint	Paper	Modelcard	Datasheet	Package	API
OLMo 7B Instruct	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	-
BLOOMZ	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓
AmberChat	✓	✓	✓	✓	✓	✓	-	-	✓	✗	-	-	✗	✓
Open Assistant	✓	✓	✓	✓	✗	✓	✓	✓	-	✗	✗	✗	✓	✓
OpenChat 3.5 7B	✓	✗	✓	✗	✓	✓	-	✓	✓	✓	-	✗	✓	-
Pythia-Chat-Base-7...	✓	✓	✓	✓	✗	✓	✓	✓	-	✗	-	-	✓	✗
Cerebras GPT 111...	-	✓	✓	✓	✓	-	✗	✓	-	✗	✗	✓	✗	✓
RedPajama-INCITE...	-	✓	✓	✓	✓	-	-	-	✗	✗	✓	✓	✗	-
dolly	✓	✓	✓	✓	✗	✓	✓	✓	-	✗	✗	✗	✓	✗
Tulu V2 DPO 70B	✓	✗	-	✓	✓	-	-	-	✓	✗	-	-	✗	✓
MPT-30B Instruct	✓	-	✓	-	✗	✓	✓	-	✗	✗	-	✗	✓	-
MPT-7B Instruct	✓	-	✓	-	✗	✓	✓	-	✗	✗	✓	✗	✓	✗
trix	✓	✓	✓	-	✗	✓	✓	-	✗	✗	✗	✗	-	✓
Vicuna 13B v 1.3	✓	-	✓	✗	✗	-	✓	✗	✓	✗	-	✗	✓	-
minChatGPT	✓	✓	✓	-	✗	✓	✓	-	✗	✗	✗	✗	✗	✓
ChatRWKV	✓	-	✓	✗	✗	✓	-	-	-	✗	✗	✗	✓	-
BELLE	✓	-	-	-	-	✗	-	✓	✓	✗	✗	-	✗	✗
WizardLM 13B v1.2	-	✗	-	✓	✓	-	-	✓	✓	✗	✗	✗	✗	✗
Airoboros L2 70B G...	-	✗	-	✓	✓	-	-	-	✗	✗	-	-	✗	✗
ChatGLM-6B	-	-	✓	✗	✗	✓	-	-	✗	-	✗	✗	✗	✓
Mistral 7B-Instruct	-	✗	✓	✗	-	✓	✗	-	-	✗	✗	✗	-	✓
WizardLM-7B	-	-	✗	✓	-	-	-	✓	✓	✗	✗	✗	✗	✗
Qwen 1.5	-	✗	✓	✗	✓	✗	-	-	✗	✗	✗	✗	-	✓
StableVicuna-13B	-	✗	-	-	-	-	-	-	-	✗	-	✗	✗	-
Falcon-40B-Instruct	✗	-	✓	-	✗	✓	✗	-	-	✗	-	✗	✗	✗
UltraLM	✗	✗	-	✓	-	✗	✗	-	✓	✗	-	-	✗	✗
Yi 34B Chat	-	✗	✓	✗	✓	-	✗	✗	✓	✗	✗	✗	✗	-
Koala 13B	✓	-	-	-	✗	-	-	-	✗	✗	✗	✗	✗	✗
Mixtral 8x7B Instruct	✗	✗	✓	✗	✓	✓	✗	-	-	✗	✗	✗	-	✗
Stable Beluga 2	✗	✗	-	✗	✓	-	✗	-	-	✗	-	✗	✗	-
Stanford Alpaca	✓	✗	-	-	-	✗	-	✓	✗	✗	✗	✗	✗	✗
Falcon-180B-chat	✗	-	-	-	-	✗	✗	-	-	✗	-	✗	✗	✗
Orca 2	✗	✗	-	✗	✓	✗	✗	-	-	✗	-	✗	✗	-
Command R+	✗	✗	✗	✓	✓	-	✗	✗	✗	✗	-	✗	✗	✗
Gemma 7B Instruct	-	✗	-	✗	-	✗	✗	-	✗	✗	✓	✗	✗	✗
LLaMA2 Chat	✗	✗	-	✗	-	✗	✗	-	-	✗	-	✗	✗	-
Nanbeige2-Chat	✓	✗	✗	✗	✓	-	✗	✗	✗	✗	✗	✗	✗	-
Llama 3 Instruct	✗	✗	-	✗	-	✗	✗	-	✗	✗	-	✗	✗	-
Solar 70B	✗	✗	-	✗	-	✗	✗	✗	✗	✗	-	✗	✗	-
Xwin-LM	✗	✗	-	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	-
ChatGPT	✗	✗	✗	✗	✗	✗	✗	✗	-	✗	✗	✗	✗	✗



TOVÁBBI FONTOS TRENDEK

- Fizikai korlátok elérése
- Internet (adatbázisok) szennyeződése
- Kell az áram
- Jogi környezet, intézmények
- Politika, Chips Act, stb.
- Továbbra sem világos üzleti modellek
- Egymás nyírása: vadnyugat



**2024-BEN MINDEN NAGYOBB CÉGNÉL
JELENTŐS MI FEJLESZTÉSEK ZAJLOTTAK
LE. ÚJ, EDDIG NEM LÁTOTT RENDSZEREK
INDULTAK EL. PÁR ÉV, ÉS MINDENHOL MI
ALAPÚ LESZ A VÁLLALATOK MŰKÖDÉSE...**

vagy mégsem?

FŐBB FÉLELMEK, VISSZATARTÓ OKOK (AVAGY ÜGYFÉL ATTITŰD 2024-BEN)

Jajj, ki ne maradjunk...

De ha meg kicsit várunk, még sokkal jobb lesz...

Nem is értjük mit csinálunk

De tök menő...

Inkább csak imitáljunk valamit,

Nem várunk semmit, viszont nem is kerül sokba



AI

AI

AI

AI

AI

AI

A

GREATUL ARTS

NETRAL LARSE

RETEL BART



RETRAL ARTS

RETRAL NETWORK

FOI

AFQ

FOOD ARTS

FOOD FOOD



MI A SZEREPÜNK?

Promptolni fogunk?

Vagy azt az ügyfeleink is tudnak?

Vagy mindenki nyugdíjba megy?

Vagy a „last mile”-on akad el az MI fejlődés?



CLEMENTINE

Magyar voicebotok 2024 riport

...majdnem

MINEK?

- Minden szállító nagyon büszke a hangasszisztensére.

Joggal?

- Vajon miért nincs még minden nagy cégnek ilyen?
- Fejlődnek ezek a botok?
- Ki használ ilyet?

MIÉRT MI?

Mert senki más nem csinálja meg...

DE HALLGASSUNK INKÁBB MEG EGY HÍVÁST!

Az automata
hogy lehet
túlterhelt?

1:13 és máris
mondhatom?

Azt hiszem ezt
alaposan kiveséztük.

De ezt más
elmondtam 😊

Ugye milyen
fontos az intro
frappáns
szövegezése?



Ez 2:53 perc volt.
Vajon lehetne ez
gördülékenyebb?



HOGY LEHET TESZTELNI?

- Tudományos tesztelés (háttérzaj pontos mérése)
- Vállalati szempontok szerinti tesztelés (mennyit spórol a cégnek, mennyi ügyintézőt tehermentesít, mennyi profitot termel)
- Ügyfél szempontú tesztelés
 - a) belső tesztelés (ha hozzáférnénk belső infókhoz, statisztikákhoz)
 - b) külső tesztelés (egyszeri ügyfél szempontjai)

HOGY LEHET TESZTELNI?

- Tudományos tesztelés (háttérzaj pontos mérése)
- Vállalati szempontok szerinti tesztelés (mennyit spórol a cégnek, mennyi ügyintézőt tehermentesít, mennyi profitot termel)
- **Ügyfél szempontú tesztelés**
 - a) belső tesztelés (ha hozzáférnénk belső infókhoz, statisztikákhoz)
 - b) **külső tesztelés (egyszeri ügyfél szempontjai)**

Voicebot	Elérhetőség		Dedikált feladat	Célcsoport	Use case
Vanda (Telekom)	telefon	1414	ügyfélszolgálat	Telekom ügyfelek	szabadszöveges ügyintézőhöz irányítás, mobilinternet bővítés, egyenleg lekérdezés
Kate (K&H)	mobil app		ügyfélszolgálat, egyszerű feladatmegoldás	K&H számlatulajdonos és mobilapp felhasználó	általános információk, utalás, limit módosítás, bankkártya letiltás
Elvira (Máv)	telefon	+3613494949	menetrend és jegyinformáció	MÁV utasok	belföldi menetrendinformáció, jegyárak
OTP Asszisztens	telefon		ügyintézőhöz irányítás, nyudíjlekérdezés	banki ügyfelek	szabadszöveges ügyintézőhöz irányítás
Vivi (Főtaxi Robotdiszpécser)	telefon	+3614222253	taxirendelés	utasok	taxirendelés felvétel

Voicebot	Elérhetőség		Dedikált feladat	Célcsoport	Use case
Vanda (Telekom)	telefon	1414	ügyfélszolgálat	Telekom ügyfelek	szabadszöveges ügyintézőhöz irányítás, mobilinternet bővítés, egyenleg lekérdezés
Kate (K&H)	mobil app		ügyfélszolgálat, egyszerű feladatmegoldás	K&H számlatulajdonos és mobilapp felhasználó	általános információk, utalás, limit módosítás, bankkártya letiltás
Elvira (Máv)	telefon	+3613494949	menetrend és jegyinformáció	MÁV utasok	belföldi menetrendinformáció, jegyárak
OTP Asszisztens	telefon		ügyintézőhöz irányítás, nyudíjlekérdezés	banki ügyfelek	szabadszöveges ügyintézőhöz irányítás
Vivi (Főtaxi Robotdiszpécser)	telefon	+3614222253	taxirendelés	utasok	taxirendelés felvétel

MIT TUDNAK “HIVATALOSAN”?

VANDA:

- kiegészítő mobilinternet bekapcsolás,
- információ folyamatban lévő megrendelésről
- hiba esetén vonalvizsgálatot végez, szükség esetén hibajegyet rögzít
- tájékoztatást ad az egyenlegről: hány lebeszélhető perc maradt, mennyi adatforgalom használható még fel, előző havi számlád végösszegéről is.
- jelszó módosítás: wifi, webtárhely, internet vagy e-mail azonosító

ELVIRA:

- belföldi menetrend információ
- tájékoztatás jegyárakról
- kedvezményekről
- későbbi korábbi vonatindulások
- kerékpár szállítás
- Sms üzenet válasz

KATE:

- utalás
- limitek beállítása (vásárlási, pénzfelvételi, online vásárlási)
- PIN-kód mutatás
- kivonat áttekintése
- infó befektetésekről, hitelekről, biztosításokról

HOGYAN MÉRHETŐ EGY VOICEBOT

(HA NINCS MÓDSZERTAN, CSINÁLJUNK!)

Kérdés: Hogyan tudjuk ezeket a szempontokat számokkal kifejezni, hogyan mérhetőek objektív és megismételhető módon?

Objektív módon csak egy része mérhető: pl. szubjektív ügyfélélmény nem.

Tesztelési körülmények:

- 1) Hangfelvételt készítünk minden tesztről, mert menet közben nehéz kiértékelni.
- 2) Olyat kérdezünk amit tudnia kell (amit ígér, hogy tud). Tényleg tudja-e őket?
- 3) Konzisztencia teszt: fix szkriptek olyan esetekre amiket tud. Ha ugyanazt kérdezzük, ugyanaz lesz-e a válasz?
- 4) Figyelembe vesszük az adott botról elérhető információkat.
- 5) Az első tesztek tapasztalatai alapján finomítjuk az irányokat, nem szivatjuk szándékosan olyan dolgokkal, ami rögtön látszik, hogy nem megy.

MÉRŐSZÁMOK

(MERT MÉRNI KELL!)

1) Konzisztencia: Ugyanazzal a kérdéssel megkeressük (fix szkript) újra és újra, ugyanazt válaszolja-e? Kijelöljük a leggyakoribb verzót és azt vesszük alapnak és annak az előfordulási aránya less a konzisztencia érték. Azonos feltételek kérdéses, hogy mennyire teljesülnek a különböző hívók miatt, ezt lehetne géppel tesztelni.

2) Önállóság: az ügyek hány százalékát tudja megoldani ügyintéző nélkül?

3) Siker arány/működési hatékonyság: összesen x hívából hányszor sikerül elintézni a kívánt dolgot? (kérdés: ez eltér-e az önállóságtól? Van-e értelme külön a kettőnek?) Siker-e ha igen, de ügyintézővel? Esetleg ketté lehet választani a következőképpen:

- a) konzisztencia teszten belüli siker arány
- b) általános siker arány

4) Visszakérdezési arány: hányszor kérdez vissza az összes kérdés válasz pár arányában? (ez lehet esetleg rugalmassági mérőszám is) - ezt melyikből kéne számolni? Lehet inkább konzisztenciából, mert Vandánál nehéz mérni, hogy hány kérdés válasz pár hangzott el, amihez arányosítanánk.

5) Szubjektív ügyfélélmény: 1-5-ig kiértékelése a beszélgetésnek, több szempont átlaga.

6) Robot beszédminőség, hangsúlyozás, természetesség (szubjektív): 1-5

7) Rendelkezésre állás: történt-e technikai hiba? Milyen "nyitvatartással" érhető el?

8) Rugalmasság: mennyire lehet változatosan fogalmazni neki?

DE HOGYAN ÉRZI EGY ÜGYFÉL?

(MINDANNYIAN ÜGYFELEK VAGYUNK!)

Szubjektív ügyfélélmény szempontok:

- Milyenek érezted a gépi beszéd folytonosságát? (1 - 5) (1 - nagyon szaggatott, 5 - folytonos)
- Milyenek érezted a gépi beszéd hangsúlyozását? (1 - 5) 1 - monoton, 2, 3, 4 - emberi, 5 - művészi
- Összességében a válaszadás gyorsasága: (1-5) 1 - lassú, 3 - közepes, 5 - gyors
- Összességében az ügyintézés gyorsaságát hogyan értékelet? (1-5) 1 - lassú, 5 - nagyon gyors
- Összességében milyen volt az élmény? (1-5) 1 - nagyon rossz, 5 - nagyon jó

A SZÁMOK PEDIG...


szempontok	VANDA	ELVIRA	KATE
1) Konzisztencia	55% (12/22)	97% (29/30)	100% (20/20)
2) Önállóság	30% (6/20)	65% (13/20)	?
3) a) Siker arány általános	60% (12/20)	95% (19/20) 75% (24/32)*	90% (9/10)
3) b) Siker arány konziszt	86% (19/22)	80% (24/30)	100% (20/20)
4) Visszakérdezési arány	30% (6/20)	23% (17/72)	50% (5/10)
5) Szubjektív ügyfélélmény (1-5)	3.35	3.92	3.88
6) Robot beszédminősége (1-5)	3.39	3.95	3.90
7) Rendelkezésre állás	100 %	100 %	< 100 %
8) Rugalmasság			
9) Ügyintézőhöz kapcsolás	60% 12/20	35% (7/20)	-

CLEMENTINE

Összegzés


Láncreakció podcast

Láncreakció 184 - A State of AI tanulmány idén is megéri a pénzét 0:00 / 50:49




Láncreakció
Clementine Data Science podcast

184. epizód
A State of AI tanulmány idén is megéri a pénzét!



Láncreakció
Clementine Data Science podcast

183. epizód
Radikális MI-optimizmus vagy kigyóolaj-szindróma?




Láncreakció
Clementine Data Science podcast

182. epizód
MI szakértők a szőnyeg legszélén

Hetedik éve adja ki az Air Street Capital ezt a kétszáz oldalas jelentést azzal a céllal, hogy összefoglalja, mi történt az elmúlt évben az MI terü...


Élmény olyan szövegeket olvasni, amelyeket igazi MI szakértők írtak és nem álltak meg a közhelyeknél. Ezen a héten két ilyen is találtunk: Darius A...

Ki lehet MI szakértő? Kinek higgyünk, kinek a tudását fogadjuk el? Tud-e érvényes állításokat tenni a az MI működéséről egy szociálpszichológus, va...




Láncreakció
Clementine Data Science podcast

181. epizód
Milyen Nobel-díjat kapjon egy MI-kutató?



Láncreakció
Clementine Data Science podcast

180. epizód
Az önkiszolgáló kasszáktól a csöves emberi tudatig



Láncreakció
Clementine Data Science podcast

179. epizód
Konferenciák édes titkai

Idén a fizikai és kémiai Nobelt is a mesterséges intelligencia egy-egy kutató-párosa kapta, úgy látszik, a svéd királyi Akadémia hallgatott az idők...

A hét hírcsokrán szerepelnek az OpenAI-től rekordszámban távozó alapítók, az online pókért megváltoztató botok, az önkiszolgáló kasszák, mint a r...

Gyuri Londonban járt az i2 felhasználói összejövetelén, Feri pedig Siófokon volt az Internet Hungary tömeges rendezvényén. Hogyan fogtak emberkeres...

PROGRAM

9.00-9.15	Érkezés, regisztráció
9.15-9.45	Hová megy ez a vonat? Körmendi György , Clementine, ügyvezető igazgató
9.50-10.15	Gépi beszédkommunikáció gyakorlati kihívásai Dr. Zainkó Csaba , egyetemi adjunktus, BME-TMIT
10.20-10.45	Hív a jövő: amikor Hanga tárcsáz Kovács-Ördög Zita , vezető elemző, Clementine
10.45-11.05	Kávészünet
11.05-11.25	Generálható-e a kreativitás? Kánai András , B2B kreatívigazgató, Meraki Marketing Kft
11:30-11.50	A jövő már itt van: AI asszisztensek szerepe a modern HMI-ekben Husztai Dániel , Data & AI Technical Sales, IBM



PROGRAM

- | | |
|--------------------|---|
| 11.55-12.10 | Képgenerálás nagyvállalati környezetben AI segítségével
Böszörményi Ádám , CX/EX Solution Sales Manager, NTT Magyarország |
| 12.15-12.30 | RAG kísérletek - LLM alapú ügyfélszolgálati asszisztens
Zsilinszki Anna , data scientist, Clementine |
| 12.30-13.30 | Ebéd |
| 13.30-13.50 | Personal SecOps: gyors receptek a személyes/eszköz kiberbiztonsághoz
Timur Khrotko , tagozatvezető, OWASP |
| 14.00-15.00 | Láncreakció podcast felvétel
Résztevők:
Pohly Ferenc, Körmendi György és vendégeik |



