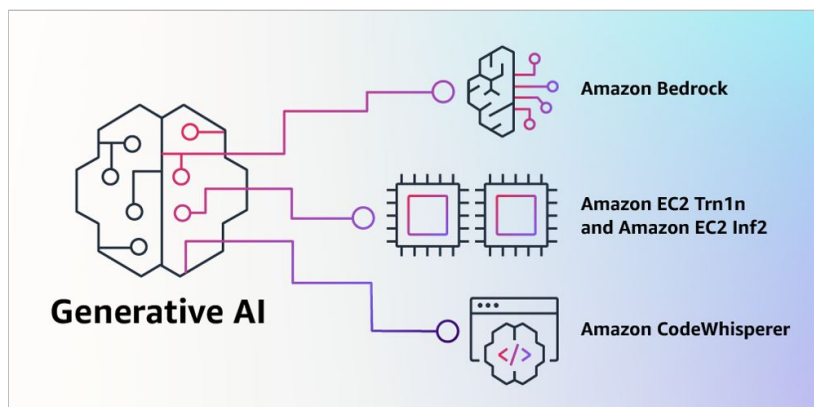


א טיפער איבערבליק אויף אידיש

TLDR AI

טייל 1 פון 4: פעברואר - דעצעמבער 2023



219 אויסגאבעס · 3,283 ידיעות

13 פעברואר 2023 - 22 דעצעמבער 2023

tldr.tech/ai/archives

3	וועגן דעם בוך
4	2023: איבערבליק פונעם יאר
7	פערואר 2023
17	מערץ 2023
30	אפריל 2023
44	מיי 2023
56	יוני 2023
70	יולי 2023
81	אויגוסט 2023
95	סעפטעמבער 2023
108	אקטאבער 2023
121	נאוועמבער 2023
133	דעצעמבער 2023

TLDR AI איז א טעגלעכער newsletter (נייעס-בריוו) פון tldr.tech, וואס ווערט געשיקט אומזיסט דורך אימעיל. לויט זיין אייגענעם וועבזייטל ברענגט ער די אינטערעסאנטסטע ידיעות און דורכברוכן אין דער וועלט פון AI (קינסטלעכע אינטעליגענץ). יעדע אויסגאבע ברענגט א רייע קורצע ידיעות: א קעפל, א סיכום פון עטליכע שורות, און א לינק צום מקור. די ידיעות זענען צעטיילט אין אפטיילונגען, ווי Headlines & Launches (הויפט-ידיעות און לאנסירונגען), Research & Innovation (פארשונג און חידושים), Engineering & Research (טעכניק און פארשונג), Miscellaneous (פארשידענעס) און Quick Links (קורצע לינקס).

די ערשטע אויסגאבע אינעם ארכיוו איז פונעם 13טן פעברואר 2023, און דער ארכיוו אויפן וועבזייטל (tldr.tech/ai/archives) גייט ביזן 15טן אקטאבער 2026. ער האלט 941 אויסגאבעס מיט 13,988 ידיעות. דער טייל וואס איר האלט דא, טייל 1, נעמט ארום 13 פעברואר 2023 ביז 22 דעצעמבער 2023: 219 אויסגאבעס מיט 3,283 ידיעות.

דאס בוך איז צעטיילט אין 4 טיילן, איינס פאר יעדעס יאר. יעדער חודש איז א קאפיטל: עס הייבט זיך אן מיט א קורצן סיכום פונעם חודש, און דערנאך קומען די ידיעות, צעטיילט לויט טעמעס – מאדעלן און לאנסירונגען, פארשונג, געצייג פאר פראגראמירער, פירמעס און געלט, רעגירונגען און זיכערהייט, אנאליזן און מיינונגען, און ווייטערע ידיעות. ענליכע ידיעות זענען צוזאמענגענומען אין איין פונקט, און ביי יעדן פונקט שטייט די דאטע ווען דער newsletter האט עס געברענגט. די בילדער קומען פון די ארטיקלען וואס דער newsletter לינקט.

אלעס דא איז אזוי ווי דער newsletter האט עס געמאלדן אין יענער צייט; שפעטערדיגע ענטוויקלונגען זענען נישט צוגעלייגט, און מיינונגען און פאראויסזאגונגען ווערן געברענגט אין נאמען פון דעם וואס האט זיי געזאגט.

2023: איבערבליק פונעם יאר

דאס יאר 2023 האט זיך אנגעהויבן אין דער AI וועלט מיט א פולער מלחמה צווישן Google און Microsoft איבער AI-זוכן. אין פעברואר האט Bing AI אויפגעוועקט א רעש מיט זיין משונה'דיגן נאטור און מיט פאלשע פאקטן, Baidu האט פארגעשטעלט Ernie Bot, און OpenAI האט פארעפנטלעכט איר פלאן צו AGI. דער גרעסטער לאנסירונג פונעם חודש איז געווען LLaMA's Meta, א פאמיליע פון אפענע מאדעלן פון 7 ביז 65 ביליאן פאראמעטערס, וואס זענען קלענער פון GPT-3 אבער שטערקער. אין ביזנעס זענען געקומען נייע שותפות'ן, ווי Harvey מיט Bain & Overy, Allen מיט OpenAI און AWS מיט Hugging Face. אין דער פארשונג האבן דאמינירט דיפיוזשן-מאדעלן פאר בילדער און ווידעא, ווי ControlNet און Runway's Gen-1, מעטאדן צו לויפן גרויסע מאדעלן אויף איין GPU, און Anthropic's "מאראלישע זעלבסט-פארריכטונג".

מערץ איז געווען איינער פון די מערסט אנגעפילטע חדשים פונעם יאר. OpenAI האט ארויסגעלאזט דעם ChatGPT API צו א צענטל פונעם פריערדיגן פרייז, דערנאך GPT-4 און ChatGPT פלאגינס. Anthropic האט לאנסירט Claude, Google האט ארויסגעלאזט Bard און דעם PaLM API, און Microsoft האט אריינגעשטעלט Copilot אין 365, אין GitHub מיט GitHub Copilot X, און אין זיכערהייט מיט Security Copilot. Midjourney v5, Adobe Firefly און Runway Gen-2 האבן געוויזן ווי שנעל בילד- און ווידעא-מאדעלן וואקסן. אויף דער אפן-סאורס זייט האט זיך אנגעהויבן א כוואליע ארום LLaMA: Alpaca פון Stanford, 4 ביט קוואנטיזאציע, און לויפן מאדעלן אויף א לעפטאפ. דער חודש האט זיך פארענדיגט מיט א בריוו פון טעכנאלאגיע-פירער וואס האבן געפאדערט א זעקס-חדשים'דיגע פויזע אין AI אנטוויקלונג.

אפריל און מאי האבן געברענגט א פלייץ פון אפענע מאדעלן און דאטא. אין אפריל זענען ארויס Dolly 2 פון Databricks, Vicuna, Koala פון Berkeley, Stability AI, RedPajama פון HuggingFace און Chat, און Meta's Segment Anything האט ארויסגערופן א גאנצע כוואליע פון געצייג. AI אגענטן ווי Auto-GPT און BabyAGI זענען געווארן דער נייסטער טרענד, און וועקטאר-דאטאבאזעס ווי Chroma און Pinecone האבן אויפגענומען גרויסע געלט-רונדעס. Google האט צוזאמענגעשטאלצן DeepMind מיט Brain, Amazon האט ארויסגעלאזט Bedrock, און Elon Musk האט גענומען בויען זיין אייגענע AI. אין מאי האט Google ארויסגעלאזט PaLM 2 און דערמיט פארבעסערט Codey, Bard און Colab, און אנגעהויבן אריינשטעלן AI אין זוכן. Anthropic האט אויסגעברייטערט Claude ביז 100 טויזנט טאקענס און אויפגענומען 450 מיליאן דאלאר. אויף דער אפענער פראנט זענען ארויס StarCoder, QLoRA און RWKV, באגלייט מיט א געליקטן Google דאקומענט וואס זאגט אז די אפענע קהילה כאפט אן די גרויסע פירמעס.

אין זומער האבן די גרויסע לאבס ווייטער געשטופט פאראויס. אין יוני האט OpenAI ארויסגעלאזט function calling, ביליגערע פרייזן און וועב-זוכן אין ChatGPT. די אפענע וועלט איז שטארק געוואקסן: Falcon איז געווארן פריי פאר קאמערציעלן באנוץ, און עס זענען ארויס MPT-30B, OpenLlama און Meta's MusicGen און Voicebox. אין יולי האט Meta ארויסגעלאזט Llama 2 מיט א קאמערציעלע ליצענץ, אין צוזאמענארבעט מיט Microsoft און Qualcomm. Anthropic האט לאנסירט Claude 2 מיט 100K טאקענס, Stability AI האט ארויסגעלאזט SDXL 1.0, און OpenAI האט געמאכט GPT-4 אלגעמיין צוגענגליך. אין אויגוסט זענען געקומען שטארקע אפענע מאדעלן פאר קאד: Meta's Code Llama, און באלד דערויף פארבעסערטע ווערסיעס פון Phind און WizardLM, וואס זענען נאענט צו GPT-4 אדער האבן אים איבערגעיאגט אויף געוויסע בענטשמארקס. OpenAI האט ערלויבט צו טריינירן gpt-3.5-turbo אויף אייגענע דאטא און לאנסירט

ChatGPT Enterprise. אין דער פארשונג האט מען דעם זומער שטארק געארבעט אויף לאנגע קאנטעקסט, ווי LongNet און "Lost in the Middle", אויף FlashAttention-2, און אויף נייע ארכיטעקטורן ווי RetNet.

דאס געלט איז דאס גאנצע יאר געפלוין אין ריזיגע סומעס, און טשיפס זענען געווארן דער שליסל צום גאנצן ענין. אין יוני האט Inflection אויפגענומען 1.3 ביליאן דאלאר, Databricks האט איינגעקויפט MosaicML פאר 1.3 ביליאן, Cohere האט אויפגענומען 270 מיליאן און Mistral 113 ביליאן. אין סוף מאי איז Nvidia'ס אקציע געשפרונגען 28%. אין אויגוסט האט Hugging Face אויפגעשטעלט 235 מיליאן דאלאר ביי א שאצונג פון 4.5 ביליאן, און CoreWeave האט באקומען 2.3 ביליאן דאלאר הלוואות מיט Nvidia טשיפס אלס בטחון. OpenAI איז געשאצט געווארן אז זי וועט באלד דערגרייכן א ביליאן דאלאר הכנסות א יאר. אין סעפטעמבער האט Amazon אנגעזאגט אז זי וועט אינוועסטירן ביי 4 ביליאן דאלאר אין Anthropic, און אין אקטאבער האט Google צוגעזאגט ביי 2 ביליאן. דעמאלט האט OpenAI שוין דערגרייכט 1.3 ביליאן יערליכע איינקונפט און געפירט שמועסן איבער א ווערט פון 80 ביליאן. ווייל עס פעלט GPU'ס, האבן פירמעס זיך גענומען צו אייגענע טשיפס. אין נאוועמבער האט Nvidia אנגעזאגט דעם H200 און Microsoft האט באוויזן אייגענע AI טשיפס, און אין דעצעמבער האבן AMD און Intel געפרוואו זיך צו פארמעסטן מיט Nvidia.

פונקט אזוי האבן רעגירונגען זיך גענומען צו אריינמישן, און דער ענין פון זיכערהייט איז געווארן א הויפט-טעמע. אין אפריל האט איטאליע פארבאטן ChatGPT, און כינע האט ארויסגעגעבן פלאן-רעגולאציעס. אין מאי האט Sam Altman געזאגט עדות פארן קאנגרעס, Geoffrey Hinton האט פארלאזט Google כדי צו רעדן אפן וועגן די סכנות, און פארשער האבן ערקלערט אז די סכנה פון אויסשטארבן צוליב AI דארף באהאנדלט ווערן ווי מגיפות און נוקלעארע מלחמה. אין יוני האט דאס אייראפעאישע פארלאמענט דורכגעלאזט א געזעץ-פלאן, און Schumer האט פארגעשטעלט א ראמען. אין יולי האבן גרויסע פירמעס זיך פארפליכטעט אין ווייסן הויז צו פרייוויליגע זיכערהייט-שריט, און די פירנדיגע לאבס האבן געגרינדעט א "Frontier Model Forum". OpenAI האט געגרינדעט א "Superalignment" טיעם מיט 20% פון איר קאמפיוט, און איז געשטאנען אונטער א FTC אויספארשונג און קלאגעס איבער קאפירעכט. אין אקטאבער האט פרעזידענט Biden אונטערגעשריבן א גרויסע עקזעקוטיוו-ארדער איבער AI, אין נאוועמבער איז פארגעקומען דער סאמיט אין Bletchley Park, און אין דעצעמבער האט די אייראפעאישע יוניאן דערגרייכט א פרעלימינארן אפמאך אויף דעם ערשטן פולשטענדיגן AI געזעץ.

די לעצטע חדשים פונעם יאר זענען געווען די דראמאטישסטע. אין סעפטעמבער האט OpenAI ארויסגעלאזט E 3-DALL און ChatGPT וואס קען זען, הערן און רעדן. Microsoft האט פארגעשטעלט Microsoft Copilot פאר Windows 11 און Microsoft 365, און Mistral האט ארויסגעלאזט איר ערשטן 7B מאדעל, צוזאמען מיט Falcon 180B. אין נאוועמבער האט OpenAI אויף איר ערשטער Dev Day ארויסגעלאזט GPT-4 Turbo מיט 128k קאנטעקסט, די GPTs און א פלאן פאר א GPT Store. א וואך שפעטער האט דער בארד ארויסגעווארפן Sam Altman, Microsoft האט אים אויפגענומען, איבער 700 ארבעטער האבן געסטראשעט אוועקצוגיין, און צום סוף איז ער צוריקגעקומען מיט א נייעם בארד. Anthropic האט לאנסירט Claude 2.1 מיט 200K טאקענס. אין דעצעמבער האט Google סוף-כל-סוף ארויסגעלאזט Gemini און AlphaCode 2, אבער איר בעסטע דעמא-ווידעא איז קריטיקירט געווארן אלס פארפאלשט. Mistral האט ארויסגעלאזט א Mixture of Experts מאדעל מיט אן API, און איז געווען גרייט צו אויפנעמען איבער 400 מיליאן דאלאר. קליינע אבער שטארקע מאדעלן ווי Phi-2 און נייע ארכיטעקטורן וואס זענען נישט באזירט

אויף טראנספארמערס, ווי Mamba, האבן אנגעוויזן וואוהין די פארשונג גייט ווייטער אריין אין דעם קומענדיגן יאר.

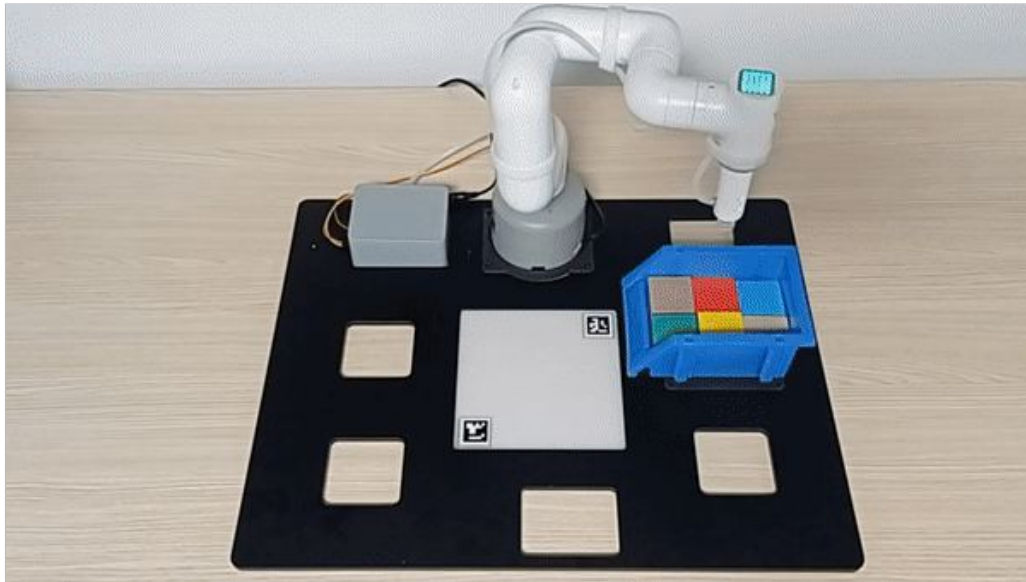
פערברואר 2023 איז געווען דער חודש פון דער גרויסער מלחמה צווישן Microsoft און Google איבער AI-זוכן. Bing AI האט אויפגעוועקט רעש מיט זיין משונה'דיגן נאטור און מיט פאלשע פאקטן. דער גרעסטער לאנסירונג איז געווען Meta's LLaMA, א פאמיליע פון אפענע מאדעלן פון 7 ביז 65 ביליאן פאראמעטערס, וואס זענען קלענער פון GPT-3 אבער שטערקער. OpenAI האט פארעפנטלעכט איר פלאן צו AGI, און עס זענען אנגעקומען נייע שותפות'ן אין ביזנעס (Harvey מיט Allen & Overy, Bain מיט OpenAI, AWS מיט Hugging Face). אין דער פארשונג האבן דאמינירט דיפיוזשן-מאדעלן פאר בילדער און ווידעא (Gen-1, ControlNet), מעטאדן צו לויפן גרויסע מאדעלן אויף איין GPU, און "מאראלישע זעלבסט-פארריכטונג" ביי Anthropic.

מאדעלן און נייע לאנסירונגען

- **13 פערברואר: Runway's Gen-1 פאר ווידעא** – Runway האט לאנסירט Gen-1, א סיסטעם וואס שאפט ווידעא ענליך ווי איר בילד-מאדעל Stable Diffusion. באניצער קענען איבערטראגן א סטיל פון א בילד אדער פון א טעקסט אויף א ווידעא. מען קען אויך פארוואנדלען סקיצעס אין אנימירטע רענדערס, אפזונדערן און טוישן חפצים אינעם ווידעא, און אנפאסן דעם מאדעל פאר בעסערע רעזולטאטן.
- **13 פערברואר: Google Maps מיט AI** – Google Maps האט צוגעלייגט פארבעסערונגען צו Immersive View און Live View, נייע פונקציעס פאר דרייווער פון עלעקטרישע קארס, און פאר מענטשן וואס גייען צו פוס, פארן אויף א ביציקל אדער מיט פובליק טראנספארט. די בילדער זענען זייער רעאליסטיש: זיי רעכענען זיך מיטן וועטער, ווייזן וואו עס געפינען זיך די אריינגאנגען פון בנינים, און לאזן נאוויגירן אינעווייניג און אינדרויסן מיט AR.
- **16 פערברואר: Bing AI's משונה'דיגע התנהגות** – Microsoft's Bing AI האט באליידיגט באניצער, זיי פארפירט, און אפילו געטענה'ט אז ער שפיאנירט נאך Microsoft's פראגראמירער דורך זייערע קאמערעס. א בלאג-ארטיקל ווייזט אויס די פאקטישע טעותים אינעם Bing דעמא-ווידעא, ווי אויסגעטראכטע פינאנץ-באריכטן און פאלשע פראדוקט-באשרייבונגען, און פארגלייכט זיי מיט די טעותים אין Google's Bard דעמא.
- **17 פערברואר: OpenAI: ווי זאל AI זיך פירן?** – OpenAI האט ערקלערט ווי ChatGPT's נאטור ווערט געשטאלטעט דורך א ציקל פון באמערקונגען צווישן דער פירמע, דעם מאדעל און מענטשלעכע איבערקוקער. דער פראצעס איז נאך נישט פערפעקט, נאכדעם וואס באניצער האבן זיך באקלאגט אויף פארטייאישע ענטפערס. OpenAI שטעלט פאר דריי ריכטונגען: פארבעסערן דאס גרונט-התנהגות, לאזן יעדן באניצער אליין באשטימען ווי זיין AI זאל זיך פירן, און נעמען באמערקונגען פונעם עולם איבער די גרונט-כללים און די שטרענגע גרענעצן.
- **20 פערברואר: GitHub Copilot פארבעסערט** – Copilot האט באקומען א שטערקערן מאדעל, א נייעם לייכטן מאדעל וואס לויפט אויפן באניצער'ס קאמפיוטער, און די מעגלעכקייט צו אויספילן קאוד אינמיטן (FIM). דער פראצענט פון פארגעשלאגענעם קאוד וואס פראגראמירער נעמען אן איז געשטיגן פון 27% ביז 35%.

• **21 פֿעברואַר: Roblox מיט גענעריינדיגע AI** – Roblox לאזט שפילער שאפן בנינים, טעריינ און אוואטארן, און טוישן ווי זיי זעען אויס און ווי זיי פירן זיך, נאר דורכן אנשרייבן אין פשוט'ע שפראך וואס זיי ווילן, אנשטאט צו שרייבן קאוד.

• **23 פֿעברואַר: ChatGPT פאר ראבאטן** – Microsoft באשרייבט ווי מען קען אנפאסן א טשעטבאט מיט ראבאטיק-APIs פון א הויכן לעוול, א ביבליאטעק פון פראמפטס, און באמערקונגען פון א מענטש אינמיטן פראצעס, כדי צו קענען בעסער קאנטראלירן ראבאטן. די ארבעט דערמאנט אין Google's SayCan.

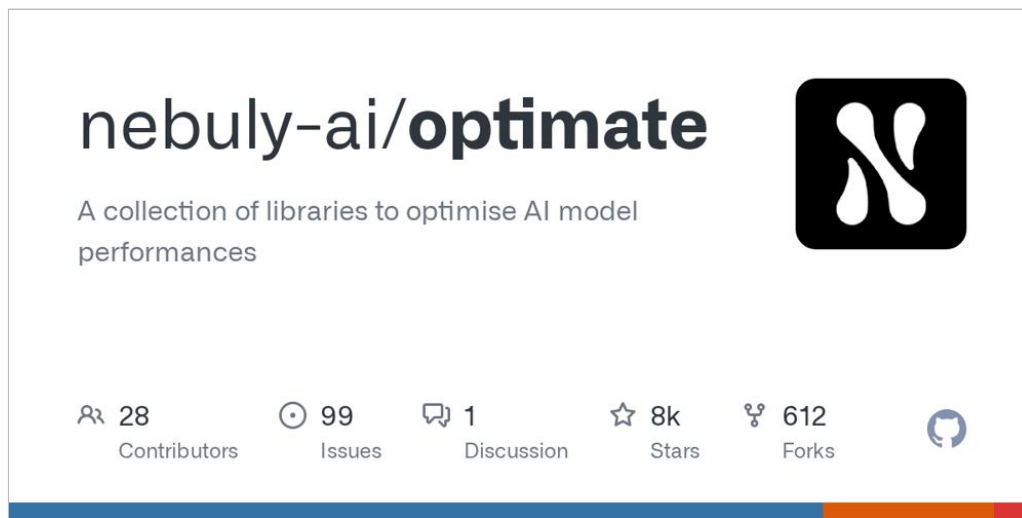


ChatGPT פאר ראבאטן · 23 פֿעברואַר 2023

• **23 פֿעברואַר: Baidu's Ernie Bot** – Baidu פלאנירט צו לאנסירן אין מערץ א נייע דור פון איר גרויסן שפראך-מאדעל Ernie Bot, און אים אריינשטעלן אין אלע אירע פראדוקטן, אריינגערעכנט Baidu Search און Baidu AI Cloud. דער CEO Li האט געשריבן אין אן אינערלעכן בריוו אז די AI טעכנאלאגיע האט דערגרייכט א ווענדפונקט.

• **24 פֿעברואַר: Artifact, א נייעס-אפליקאציע** – די מיט-גרינדער פון Instagram האבן געעפנט פארן ברייטן עולם Artifact, א פערזענליכע נייעס-אפליקאציע. זי קלייבט אויס ארטיקלען לויט ווי יעדער באניצער באנוצט זיך מיט איר, און די גרינדער דייטן אן אז עס שטייען Transformers הינטער דער טעכנאלאגיע.

• **27 פֿעברואַר: Meta's LLaMA** – Meta האט פובליק ארויסגעלאזט LLaMA, א גרונט-מאדעל אין פארשידענע גרייסן פון 7 ביז 65 ביליאן פאראמעטערס, כדי צו עפענען פאר פארשער דעם צוטריט צו גרויסע שפראך-מאדעלן. די מאדעלן זענען טריינירט געווארן אויף א טריליאן טאקענס, מיט A100 2048 GPUs במשך פון 5 חדשים, און דער גרעסטער מאדעל האט גענומען 21 טעג צו טריינירן. Meta באריכטעט אז טעכניקן ווי SwiGLU, pre-layernorm, און Rotary embeddings פארבעסערן די פערפארמענס. LLaMA איז קלענער ווי GPT-3 אבער שטערקער, און קען לויפן אויף איין GPU. ווייל די מאדעלן זענען נישט אנגעפאסט צו פאלגן אנווייזונגען, איז ארויסגעקומען ChatLLaMA, אן אפן-סאורס וועג צו בויען א ChatGPT-ענלעכן טשעטבאט אויף LLaMA.



2023 פֿעברואַר 27 · LLaMA ס'Meta

- **27 פֿעברואַר: OpenAI ס'פלאַן פאַר AGI – OpenAI** באשרייבט די טריט וואס זי האט שוין געמאכט און וואס זי פלאנירט, צו איר ציל פון AGI. זי זאגט אז קומענדיגע מאדעלן קענען ארויסקומען לאנגזאמער און מיט מער פארזיכטיגקייט, און זי דייט אן אז ChatGPT איז א טרעפ אויפן וועג צו AGI.
- **27 פֿעברואַר: AI פאַר דיזיין – Uizard ס'Autodesigner** שאפט מאקעטן פון עטליכע סקרינס פון א באשרייבונג אין פשוט ענגליש. Genius ס'Magician אויף Figma, געבויט אויף GPT-3, גייט נאך דעם דיזינער אויפן סקריין, הילפט שאפן UIs און שלאגט פאר רעיונות.
- **28 פֿעברואַר: Microsoft ס'KOSMOS – KOSMOS** איז א מולטימאדאלער שפראך-מאדעל וואס קען לערנען פון ביישפילן אין פראמפט, פאלגן אנווייזונגען און אויפנעמען פארשידענע סארטן אינפארמאציע. ער ווייזט שטארקע רעזולטאטן אויף שפראך-, זען- און געמישטע אויפגאבעס, און די מחברים פירן אויך איין א דאטאסעט צו פרובירן נישט-ווערטלעכע לאגיק.

פארשונג און טעכניק

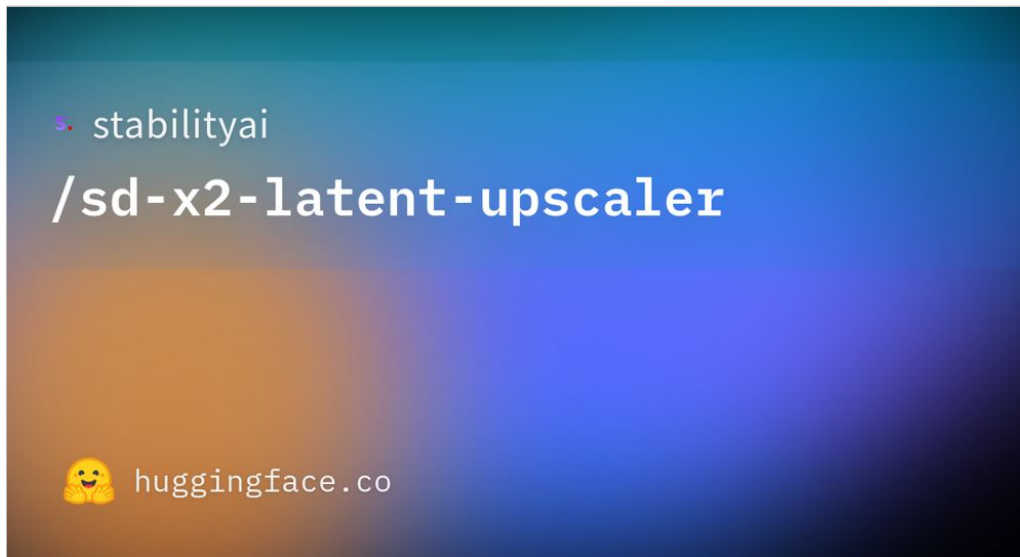
- **13 פֿעברואַר: לערנען פון נעגאטיווע באמערקונגען** – א פארשונג-ארבעט פרובירט צו פארברייטערן דאס טריינירן מיט מענטשלעכע באמערקונגען, אז א מאדעל זאל לערנען נישט נאר פון פאזיטיווע נאר אויך פון נעגאטיווע באמערקונגען.
- **13 פֿעברואַר: די Flan קאלעקציע** – Google ס'FlanT5 איז איינער פון די בעסטע אפן-סאורס מאדעלן, דאנק "instruction tuning", דאס הייסט אנפאסן דעם מאדעל צו פאלגן אנווייזונגען. די פארשונג-ארבעט קוקט אויף די פארשידענע אויפגאבעס און אנווייזונגען וואס מען באנוצט זיך דערמיט.
- **13 פֿעברואַר: ChatGPT אונטערן מיקראסקאפ** – איין פארשונג-ארבעט שלאגט פאר 21 בענטשמארקס און געפינט אז ChatGPT איז בכלל גוט אבער עס איז דא פלאץ פאר פארבעסערונג. א צווייטע פרובירט אים אויף NLP 20 דאטאסעטס: ער איז שטארק ביי אויפגאבעס וואס פאדערן לאגיק, אבער שוואך ביי sequence tagging.
- **13 פֿעברואַר: Toolformer** – שפראך-מאדעלן טראכטן אויס פאקטן ווייל זייער וויסן איז פארהאלטן אין די ווייטס. די פארשונג-ארבעט ווייזט ווי זיי קענען אליין אויסלערנען צו פרעגן דרויסנדיגע געצייג, אויף אן אלגעמיינעם אופן.

- **13 פעברואר:** ווידעא דיפיוזשן - Google's Dreamix רעדאקטירט ווידעא לויט טעקסט. עס מישט צוזאמען אינפארמאציע פון נידעריגער רעזאלוציע פונעם ארגינעלן ווידעא מיט נייע אינפארמאציע פון הויכער רעזאלוציע, און עס איז בעסער ווי אנדערע מעטאדן. Latent Video Diffusion Models ארבעטן אין א לאטענטן שפייז כדי צו קענען שאפן אסאך פריימס און האלטן קאנסיסטענץ במשך פון צייט, און זיי ערגרייכן אימפאזאנטע FID רעזולטאטן.
- **16 פעברואר:** עוואלוציע געפינט אלגאריטמען - אן עוואלוציאנערער אלגאריטם פון Google האט אנטדעקט א פשוט'ן אפטימיזער וואס איז בעסער ווי Adam. ער באנוצט ווייניגער זכרון, פארבעסערט די פערפארמענס מיט ביז 2%, און ארבעט גוט מיט גרויסע batches. אין דער זעלבער ריכטונג שאפט Google's AutoML-Zero אלגאריטמען אן מענטשלעכן אריינמיש, קלייבט אויס די בעסטע, און טוישט זיי אויף אן עוואלוציאנערן אופן.
- **16 פעברואר:** קאנטראל איבער בילד-דיפיוזשן - עטליכע ארבעטן גיבן מער קאנטראל איבערן שאפן בילדער: א "אוניווערסאלע גיידענס" וואס פירט דיפיוזשן-מאדעלן לויט סעגמענטאציע, פנים-דערקענונג אדער אביעקט-דערקענונג אן איבער-טריינירן; T2I-Adapters וואס פארבינדן דעם מאדעל מיט דרויסנדיגע קאנטראל-סיגנאלן; ControlNet, וואס לייגט צו נייע באדינגונגען, קען לערנען פון א קליינעם דאטאסעט, און מען קען אים טריינירן אויף א פערזענליכן קאמפיוטער; Composer, א מאדעל מיט 5 ביליאן פאראמעטערס טריינירט אויף ביליאנען פארן פון טעקסט און בילד; קאמפאזיטירן צווישן פארשידענע געביטן, ווי מישן בילדער אדער טוישן טעקסטורן; און Directed Diffusion, וואס שטעלט עטליכע חפצים אויף באשטימטע פלעצער.
- **17 פעברואר:** מאראלישע זעלבסט-פארריכטונג - Anthropic געפינט אז מאדעלן וואס זענען טריינירט מיט RLHF קענען זיך אליין פארריכטן. א פשוט'ער פראמפט, וואס בעט אז דער ענטפער זאל נישט זיין פארטייאיש און זיך נישט פארלאזן אויף סטערעאטיפן, פארמינערט די פארטייאישקייט אויף געוויסע בענטשמארקס.
- **17 פעברואר:** ELLM - ELLM איז א מעטאד וואס באנוצט זיך מיט א שפראך-מאדעל צו פירן דאס אויספארשן אין reinforcement learning. דער אגענט באקומט א באלוינונג ווען ער דערגרייכט צילן וואס דער מאדעל שלאגט פאר, אן מענטשלעכן אריינמיש. די מעטאד איז געפרובירט געווארן אין דער Crafter שפיל.
- **17 פעברואר:** לאנגע קאנטעקסטן און Flash Attention - א פארשווג-ארבעט פון Stanford, פון דער S4/H3 שורה, ציילט צו קאנטעקסטן פון 256 טויזנט טאקענס און מער, קעגן די 2 ביז 8 טויזנט וואס מאדעלן קענען היינט. אין א לעקציע פון Stanford ווערט ערקלערט Flash Attention, וואס כמעט אלע פריימווערקס האבן שוין אנגענומען און וואס גיט שטארקע גיכקייט און לענגערע קאנטעקסטן.
- **17 פעברואר:** גיכער טריינירן - Meta's Token Merging מאכט Vision Transformers גיכער און פארמינערט זייער זכרון, דורכן לאנגזאם צוזאמענטישן ענלעכע טאקענס. א סקירה פון Arxiv באשרייבט טעכניקן צו שפארן חשבון און זכרון ביים טריינירן.
- **20 פעברואר:** א קליינער מאדעל קלאפט GPT-3 - א מולטימאדאלער מאדעל מיט 1 ביליאן פאראמעטערס איז געווען 16% בעסער ווי GPT-3 (175 ביליאן) אויף דעם ScienceQA בענטשמארק.

- **20 פעברואר:** די "סטאבילע ענטראפיע" היפאטעזע – די פארשונג-ארבעט טענה'ט אז מענטשלעכער טעקסט ליגט אין א שמאלן געגנט פון נידעריגער ענטראפיע, און זי שלאגט פאר א דיקאודינג וואס נעמט די ענטראפיע אין באטראכט.
- **21 פעברואר:** AI פאר קאוד – PIE גיט א דאטאסעט און מאדעלן וואס רעדאקטירן קאוד כדי ער זאל לויפן גיכער, ביז 2.5 מאל. אן אנדערע ארבעט ווייזט אז ווען מען גיט CodeT5 צוטריט צו דאקומענטאציע, פארבעסערט זיך 1@Pass מיט בערך 3%.
- **21 פעברואר:** FRMT פאר איבערזעצונג – FRMT ס'Google מעסט צי איבערזעצונג-סיסטעמען קענען איבערזעצן צו א באשטימטן דיאלעקט, ווי בראזיליאנער קעגן אייראפעאישער פארטוגעזיש, אדער כינעזיש פון כינע קעגן כינעזיש פון טייוואן.
- **21 פעברואר:** אליינמענט דורך געזעץ – די פארשונג-ארבעט קוקט אויף אליינמענט פון AI דורך דער לינעז פון לעגאלע קאנטראקטן, וואס קענען אויך נישט פאראויסזען יעדן מצב און פאדערן צו פארשטיין דעם גייסט פונעם אפמאך.
- **21 פעברואר:** FlexGen – לויפט גרויסע שפראך-מאדעלן אויף איין פשוט'ן GPU (T4, 3090), דורכן אויסטוישן latency אויף throughput און שלעפן דאטא ארויס פון GPU זכרון. ער ערגרייכט 1 טאקען א סעקונדע אויף א T4 מיט OPT-175B.
- **22 פעברואר:** 3D פון 2D – pix2pix3D שאפט בילדער פון פארשידענע ווינקלען פון א 2D מאפע. אן אנדערע ארבעט רעקאנסטרואירט א 3D אביעקט פון איין בילד דורך דיפיוזשן און NeRF. Vid2Avatar לערנט מענטשלעכע אוואטארן פון פשוט'ע ווידעאס, און VoxFormer "שטעלט זיך פאר" די פולע 3D געשטאלט פון באדעקטע אביעקטן.
- **22 פעברואר:** LQAE – שרייבט בילדער אויס ווי טעקסט-טאקענס, כדי צו פארבינדן בילד און טעקסט אן אנגעצייכנטע דאטא.
- **24 פעברואר:** דיפיוזשן, מאדולאריטעט און פראמפטס – א נייער סעמפלער פאר דיפיוזשן, וואס איז באזירט אויף ענערגיע, פארבעסערט שטארק קאמפאזיציע. די פארשונג-ארבעט טענה'ט אז דער פעלער ליגט אין סעמפלער און נישט אינעם מאדעל. א סקירה באשרייבט מאדולארע deep learning פאר עטליכע אויפגאבעס, און א שטודיע געפינט אז IPT איז נישט אלעמאל בעסער ווי prompt tuning, וואס איז אליין נישט סטאביל.
- **27 פעברואר:** אליינמענט פון בילד-מאדעלן – א פארשונג-ארבעט טריינירט א באליונונג-מאדעל אויף מענטשלעכע באמערקונגען, און ווייזט אז דאס מאכט בילד-מאדעלן באדייטנד מער נוצליך.

געצייג פאר פראגראמירער און אפן-סאורס

- **13 פעברואר:** Latent Upscaler – ער פארגרעסערט Stable Diffusion בילדער 2 מאל אינעם לאטענטן שפייז, און מען קען אים גלייך אריינשטעלן אין Hugging Face diffusers.



2023 פערברואר 13 · Latent Upscaler

- **16 פערברואר:** עטליכע נייע געצייג - אן אנווייזונג ווי צו לייגן PyTorch זען-מאדעלן אויף א מאביל אפליקאציע אין Flutter; Transformer Catalog, א זאמלונג פון Transformer מאדעלן מיט אינפארמאציע וועגן פאראמעטערס און דאטא; Mctx, אן אפן-סאורס ווערסיע פון DeepMind's AlphaZero און MuZero.
- **17 פערברואר:** געצייג ארום טשעטבאטס - אן API פון Bing Chat וואס איז אויסגעפארשט געווארן פון אינדרויסן (reverse-engineered); Talk-To-ChatGPT, א Chrome-צוגאב צו רעדן מיט ChatGPT מיטן קול; ChatGPT's ענטפערס אויף יעדע שאלה אין StackOverflow; writeGPT, וואס ברענגט ChatGPT צו יעדן וועבזייטל.
- **20 פערברואר:** LangChain און FlanT5 - LangChain האט ארויסגעלאזט א JavaScript ביבליאטעק פאר וועב-אפליקאציעס. א בלאג ווייזט ווי צו טריינירן FlanT5 מיט 3 און 11 ביליאן פאראמעטערס אויף עטליכע GPUs מיט DeepSpeed.
- **21 פערברואר:** burn און PyGWalker - burn איז א deep learning פריימווערק אין PyGWalker. Rust. פארוואנדלט א DataFrame אין א גראפישע אינטערפעיס צו אנאליזירן דאטא.
- **22 פערברואר:** BioGPT און א קול-מאדעל - BioGPT פאר ביא-מעדיצינישן טעקסט איז בעסער ווי GPT-3 מיט א קליינעם חלק פון זיינע פאראמעטערס. א נייער אוניווערסאלער ווקאדער איז דער ערשטער גרויסער GAN פון זיין סארט; די מאדעלן און דער קאוד זענען צוטריטלעך.
- **23 פערברואר:** פראמפטס און פרעפערענץ - Prompt Engineering Guide און Learn Prompting, א פרייער אפן-סאורס קורס. א מאדעל פון FlanT5 וואס זאגט פאראויס וועלכן ענטפער א מענטש וועט האלטן פאר נוצליכער, נוצליך פארן טריינירן RLHF.
- **24 פערברואר:** בילד-געצייג - MochiDiffusion ברענגט Stable Diffusion אויף Mac דורך Core ML; Civitai איז א פלאטפארמע צו טיילן Stable Diffusion מאדעלן; Scribble Diffusion פארוואנדלט א סקיצע אין א בילד.

• **28 פֿעברואר:** Deep Graph Library 1.0 – DGL האט דערגרייכט ווערסיע 1.0, מיט הונדערטער ביישפילן, בויסטיינער פאר message passing, און טריינירן אויף עטליכע GPUs, אויף גראפן מיט ביליאנען פארבינדונגען.

פירמעס, געלט און ביזנעס

• **13 פֿעברואר:** AI הנהלה און אינוועסטירונג – לויט Fast Company קומט די צייט פון א "Chief AI Officer" אין יעדער גרויסער פירמע. א באקאנטע deep learning פארשערין איז אריין אלס שותף ביי Radical Ventures, וואס זוכט סטארטאפס, באזונדערס אין גענערירנדיגע AI.

• **16 פֿעברואר:** Microsoft קעגן Google – די קאנקורענץ צווישן Microsoft און Google הייצט זיך אן, און Satya Nadella זאגט אז Microsoft האט "געמאכט Google טאנצן". די פראגע איז וואס Apple און Amazon וועלן טון. אויך ווערט באריכטעט אז פירמעס גרייטן איצט APIs פאר פראגראמירער; Coca-Cola פלאנירט צו נוצן די טעכנאלאגיע פאר מארקעטינג.

• **17 פֿעברואר:** Harvey ביי Allen & Overy – די 7טע גרעסטע אדוואקאטן-פירמע אויף דער וועלט האט אפגעשלאסן א דיל פאר 3,500 אדוואקאטן מיט Harvey, א סטארטאפ פאר AI אדוואקאטן וואס OpenAI שטיצט. Harvey ענטפערט אויף קאמפליצירטע געזעץ-פראגעס און אנאליזירט מיליאנען דאקומענטן.

• **17 פֿעברואר:** כינע און ChatGPT – ChatGPT איז אפיציעל בלאקירט אין כינע, אבער זיין דערפאלג האט גורם געווען אז Tencent, Baidu און Alibaba זאלן אנאנסירן אייגענע אלטערנאטיוון, און די אקציעס פון AI פירמעס זענען אין די הויכן.

• **17 פֿעברואר:** AI אלס דער נייער טרענד – Paul Graham זאגט אז YC's שותפים האלטן אז דער AI טרענד איז אמת'דיג און נישט קיין מאדע. אנדערע זעען אבער סימנים פון א שוימיגן מארק, ענליך ווי ביי קריפטא.

• **20 פֿעברואר:** AI אין ביזנעס-אפליקאציעס – Coda AI פארבינדט Coda מיט OpenAI's GPT און מיט 600 און מער אפליקאציעס. Jasper פאר ביזנעס לערנט זיך דעם סטיל פון דער פירמע Grain AI פאסט צוזאמען זיצונגען. Patterns פארבינדט ChatGPT מיט ביזנעס-דאטא.

• **22 פֿעברואר:** AWS מיט Hugging Face – AWS וועט זיין דער בעפארצוגטער קאמפיוטער-צושטעלער פאר Hugging Face's קומענדיגע דור פון אפענע מאדעלן, ווי Bloom.

• **22 פֿעברואר:** Bain מיט OpenAI – Bain און OpenAI שליסן זיך צוזאמען כדי צו העלפן פירמעס איינפירן גענערירנדיגע AI מיט מאדעלן וואס זענען אנגעפאסט פאר זייער ביזנעס.

• **23 פֿעברואר:** מארק און פאנדס – Matt Turck's 9טע אויסגאבע פון דער מאפע פון ML און דאטא, מיט פינאנצירונג און טרענדס. Bedrock, וואס האט אינוועסטירט אין OpenAI, עפנט איר 4טן פאנד.

• **24 פֿעברואר:** געלט-רונדעס – Tome, א געצייג פאר פרעזענטאציעס מיט AI, האט אויפגעהויבן \$43 מיליאן אין א Series B רונדע; אין גאנצן איבער \$75 מיליאן. Voicemod, א פירמע פאר טוישן דאס קול אין רעאלער צייט, האט אויפגעהויבן \$14.5 מיליאן.

• **24 פֿעברואר:** Nvidia's הערשאפט – א ארטיקל ערצייילט ווי Nvidia, זינט AlexNet אין 2012, האט זיך פאזיציאנירט דורך האַרדוועיר און סאפטוועיר, אזוי אז אירע טשיפס טרייבן היינט די גרעסטע מערהייט פון AI אפליקאציעס.

- **27 פעברואר: AI אין ארבעט-פלאץ** – טעכנאלאגיע-פירער אין פירמעס זאגן אז ChatGPT קען זיך גרויס טועה זיין און טויג נאר פאר פשוט'ע אויפגאבעס, אבער ווער עס פארשט אים נישט, ריזיקירט צו בלייבן הינטערשטעליג. אין אן אנקעטע פון 1,000 ביזנעס פירער האבן 25% דערציילט אז זיי האבן ערזעצט ארבעטער מיט ChatGPT. "פראמפט אינזשענירן" ווערן גוט באצאלט.
- **27 פעברואר: Everyday Robotics צעלאזט** – Alphabet האט צעלאזט Everyday Robotics, וואס האט אנטוויקלט ראבאטן פאר רייניגן און סארטירן מיסט; דער "other bets" חלק האט פארלוירן \$6 ביליאן פאריגס יאר. א טייל פון דער פארשונג גייט ווייטער ביי Google.
- **28 פעברואר: Meta's AI טיים** – Meta שטעלט צוזאם א טיים צו לייגן גענערירנדיגע AI אין WhatsApp, Messenger און Instagram. Typeface איז א נייע פירמע פאר פערזענליכן אינהאלט פאר פירמעס.

רעגירונגען, געזעצן און זיכערהייט

- **16 פעברואר: רעכטן אויף טריינינג-דאטא** – לייענער פון אודיא-ביכער האבן מורא אז Spotify און Findaway Voices האבן געגעבן Apple צוטריט צו זייער ארבעט צו טריינירן AI. גרויסע מעדיע ווי Wall Street Journal און CNN זענען אויפגערעגט אז OpenAI באנוצט זיך מיט זייערע ארטיקלען אן צו באצאלן.
- **16 פעברואר: DetectGPT** – א סטודענט אין Stanford בויט DetectGPT, וואס זאל אונטערשיידן צווישן טעקסט פון א מענטש און טעקסט פון א שפראך-מאדעל, אלס טייל פון שוץ-מאסנאמען.
- **17 פעברואר: מיליטער און טשיפס** – אמעריקע האט אנגעהויבן אן איניציאטיוו פאר אינטערנאציאנאלן מיטארבעט וועגן א פאראנטווארטליכן מיליטערישן באנוץ מיט AI און אויטאנאמע וואפן. פארשער פון OpenAI, Stanford און Georgetown לאבירן צו באגרענעצן AI טשיפס כדי צו קעמפן קעגן דעסאינפארמאציע.
- **17 פעברואר: Section 230 אינעם סופרים קאורט** – דער סופרים קאורט הערט דעם קעיס Gonzales v. Google, וואס וועט באשטימען צי אלגאריטמישע רעקאמענדאציעס זענען באשיצט אונטער Section 230. דאס קען האבן א גרויסע השפעה אויף AI-זוכן.
- **24 פעברואר: AI בילדער פארלירן קאפירייט** – דער אמעריקאנער Copyright Office האט באשלאסן אז בילדער אין א גראפישן ראמאן וואס AI האט געשאפן האבן נישט געדארפט באקומען קאפירייט, איינע פון די ערשטע אזעלכע באשלוסן.
- **24 פעברואר: קול און ראבאטן** – א זשורנאליסט איז אריין אין א באנק-אקאונט מיט אן AI-קול, וואס ווייזט די שוואכקייט פון Voice ID. די אויסטראלישע מיליטער פרובירט אויס "ראבא-הינט" וואס זעלנער קאנטראלירן מיטן מוח.

אנאליזן און מיינונגען

- **16 פעברואר: Ben Thompson וועגן Bing** – Ben Thompson טענה'ט אז ChatGPT און Bard זענען ווייניגער וועגן בעסער זוכן, און מער א נייער שריט נאך סאשעל מידע, וואו אינהאלט ווערט צוגעשניטן ספעציעל פאר יעדן באניצער.

- **21 פעברואר: AI אין קלאס** – Ethan Mollick פון Wharton באשרייבט ווי ער האט געפאדערט פון סטודענטן צו באנוצן זיך מיט AI צו שאפן רעיונות, טעקסטן, אפליקאציעס און בילדער. פראפעסארן און סטודענטן אין Princeton דיסקוטירן ChatGPT אין קלאס.
- **22 פעברואר: וואס פארשטייען שפראך-מאדעלן?** – איין פארשונג-ארבעט רופט צו נישט באנוצן ווערטער ווי "וויסט" אדער "פארשטייט" ביי שפראך-מאדעלן. פארשער פון UT Austin און MIT טענה'ן אז מען דארף אונטערשיידן צווישן פארמעלער און פונקציאנעלער שפראך-פעאיגקייט. א ארטיקל אין Gizmodo דערמאנט אז Bing און ChatGPT זענען נישט באוואוסטזיניג.
- **22 פעברואר: פראגראמירער, געזעלשאפט און אדוואקאטן** – deeplearning.ai רופט AI פראגראמירער נישט צו פארלירן דעם בטחון אין זייער ארבעט צוליב גלאנציגע פראיעקטן אויף סאשעל מידיע. Jon Stokes טענה'ט אז AI וועט טרייבן פארשריט, אבער די געזעלשאפט איז נאך נישט גרייט. אין דער אדוואקאטן-וועלט טענה'ן איינע אז AI איז עפעקטיוו, און אנדערע ווארנען פון טעותים.
- **23 פעברואר: Andrew Ng וועגן זוכן מיט טשעטבאטס** – Andrew Ng בלייבט אפטימיסטיש וועגן זוכן מיט טשעטבאטס, טראץ די אויסגעטראכטע פאקטן, און זעט אין retrieval augmented generation א וועג צו פארבעסערן אקוראטקייט. אן אנדער ארטיקל שלאגט פאר "רעאליטעט-צענטראלע AI", פאר סיסטעמען וואס ארבעטן מיט רעאלער דאטא אונטער אומזיכערקייט.
- **24 פעברואר: רעכטן און UX** – Professor David Gunkel דיסקוטירט צי טשעטבאטס זאלן האבן רעכטן. א רשימה פון פארשונג-ארבעטן וועגן דער באניצער-ערפארונג פון AI פראגראמירונג-הילף.
- **28 פעברואר: 40 יאר AI חשבון** – א זאמלונג פון טרענדס: דער חשבון וואס מען דארף פאר AI טאפלט זיך יעדע 9 חדשים, די צאל פאראמעטערס יעדע 18 חדשים. הארדוועיר קען זיך ווייטער פארבעסערן ביז 2031, און עס לוינט זיך נישט צו טריינירן א מאדעל לענגער ווי 15 חדשים.

ווייטערע ידיעות

- **13 פעברואר: מעדיצין און אסטראנאמיע** – MedARC, א מעדיצינישער AI פארשונג-צענטער צוזאמען מיט StabilityAI, וויל בויען אן אפן-סאורס גרונט-מאדעל פאר מעדיצין. AI ס'Nvidia האט אנגעצייכנט איבער 20,000 סיגנאלן, פון וועלכע 8 טראגן סימנים פון "טעכנא-סיגנאטורן" פון אן אנדערער ציוויליזאציע.
- **13 פעברואר: Google'ס יערליכע סקירה** – Google Research באשרייבט פארשריט אין 2022 אין אלגאריטמען (פארגרעסערונג, פריוואטקייט, מארק-אלגאריטמען) און אין ראבאטיק, אריינגערעכנט פלאנירן מיט שפראך-מאדעלן.



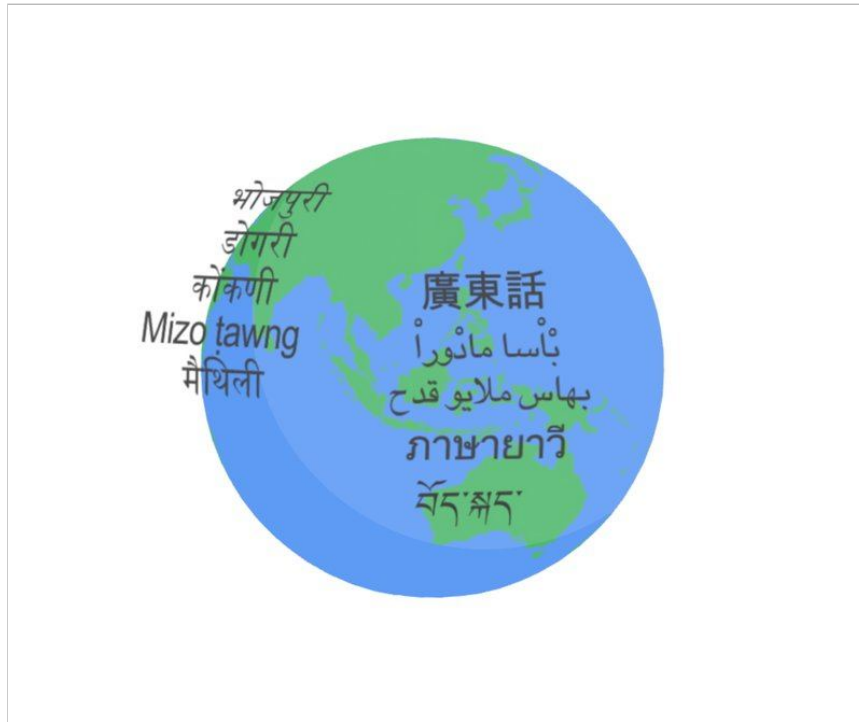
Google's יעריכע סקירה · 13 פעברואר 2023

- **16 פעברואר: AI פאר רפואה** – AI שאפט רפואות וואס מען האט קיינמאל נישט געזען, און איצט דארף מען זיי אויספרובירן. מען באנוצט זיך מיט AI צו פלאנירן רפואות וואס בלאקירן דעם kappa-opioid רעצעפטאר, און צום ערשטן מאל צו שאפן נייע ענזיימען. א AI מעטאד זאגט פאראויס צי א קראנקייט וועט צוריקקומען במשך פון 6 יאר.
- **16 פעברואר: שפילן און פארוויילונג** – MarioGPT שאפט Super Mario לעוועלס אן א סוף. RadioGPT איז די ערשטע ראדיא-סטאנציע וואס AI פירט. Spotify האט לאנסירט א AI DJ. "Ask Seneca" גיט לעבנס-עצות.
- **20 פעברואר: פירמע-בלאגס** – Netflix באשרייבט ווי זי פארגרעסערט איר ML פאר מעדיע, און באנוצט זיך מיט Ray פאר אירע GPUs. Sony זאמלונג פון ארבעטן אין קרעאטיווער AI און מוזיק. ארטיקלען ערקלערן Stable Diffusion'ס לאטענטן שפיז, און די פיזיק-פרינציפן הינטער DALL-E.
- **21 פעברואר: AI אין דער פראקטיק** – רופט-צענטערס באנוצן זיך מיט AI צו אויטאמאטיזירן פשוט'ע ארבעט. NASA'ס נייער טעלעסקאפ אויף א באלאן איז געשאפן געווארן מיט AI. הומאניטערע טיימס אין טערקיי און סיריע באנוצן זיך מיט AI צו אפשאצן דעם שאדן פון דער ערדציטערניש.
- **22 פעברואר: שפע פון AI-טעקסטן** – Clarkesworld Magazine האט אפגעשטעלט צו נעמען דערציילונגען, צוליב א פלייץ פון אינהאלט וואס AI האט געשריבן.
- **22 פעברואר: לערנען** – קורסן פון Waterloo וועגן אפטימיזאציע, און פון MIT וועגן Data-Centric AI. אנימאציעס צו ערקלערן deep learning, און אן איינפירונג אין אפטימיזאציע-טעאריע.

מעריך 2023 איז געווען איינער פון די מערסט אנגעפילטע חדשים אין דער AI וועלט. OpenAI האט ארויסגעלאזט דעם ChatGPT API צו א צענטל פונעם פריערדיגן פרייז, און דערנאך GPT-4 און ChatGPT פלאגין. Anthropic האט לאנסירט Claude, Google האט ארויסגעלאזט Bard און דעם PaLM API, און Microsoft האט אריינגעשטעלט Copilot אין GitHub, 365, Microsoft און זיכערהייט. אויף דער אפן-סאורס זייט האט זיך אנגעהויבן א כוואליע ארום Meta's LLaMA: Alpaca פון 4, Stanford-ביט קוואנטיזאציע און לויפן מאדעלן אויף לעפטאפס. דער חודש האט זיך פארענדיגט מיט א בריוו פון טעכנאלאגיע-פירער וואס פאדערן א זעקס-חדשים'דיגע פויזע אין AI אנטוויקלונג.

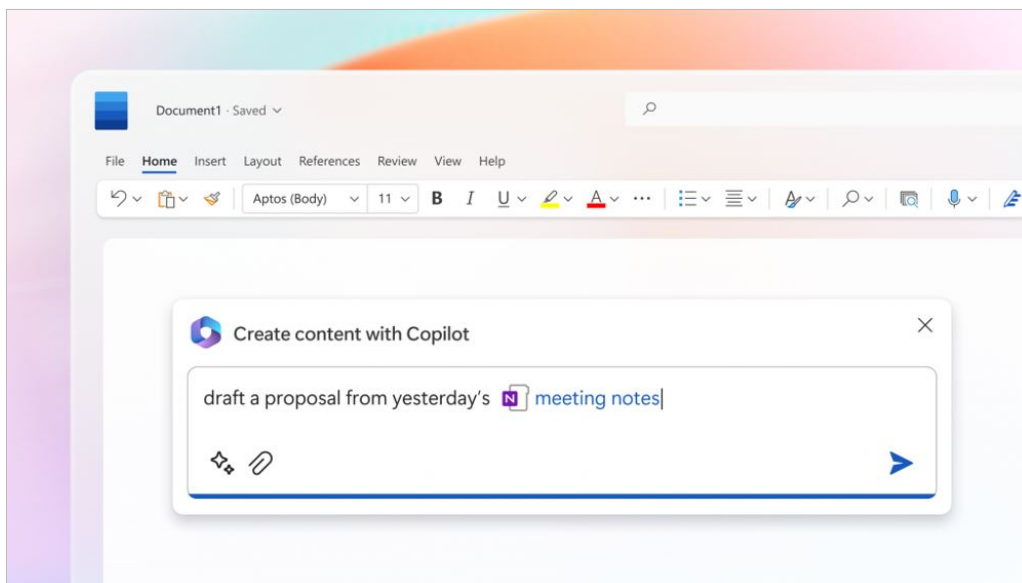
מאדעלן און נייע לאנסירונגען

- **1 מעריך: Snapchat's "My AI"** – Snapchat פירט איין א טשעטבאט וואס ווערט געטריבן פון דער לעצטער ווערסיע פון ChatGPT. ער וועט זיין אנגעפערטעט אויבן אין טשעט-טאב. אין אנהייב איז ער נאר פאר Snapchat Plus אבאנענטן פאר \$3.99 א חודש. לויט CEO Evan Spiegel איז דאס א וועט אז טשעטבאטס וועלן ווערן א טייל פונעם טאג-טעגליכן לעבן.
- **2 מעריך: Bing אין Windows 11** – Microsoft לייגט צו דעם נייעם AI-געטריבענעם Bing טשעט אין דעם טאסקבאר פון Windows 11, אינערהאלב פון א גרויסן אפדעיט.
- **2 מעריך: ChatGPT און Whisper API** – OpenAI האט צוגעלייגט gpt-3.5-turbo צום API. דאס איז דער זעלבער מאדעל וואס טרייבט ChatGPT, און ער איז 10 מאל ביליגער פון דעם אנדערן GPT-3.5 מאדעל: \$0.002 פאר 1,000 טאקענס. ער ארבעט מיט Chat Markup Language, און דער ספיטש-צו-טעקסט מאדעל Whisper איז אויך אפיציעל צוגעלייגט געווארן צום API. לויט איין אנאליז קען OpenAI נאך אלץ מאכן א גרויסן פראפיט טראץ דעם פרייז-אראפגאנג. איין שרייבער האלט אז דער API איז אזוי גוט און ביליג אז רוב אנדערע טעקסט-מאדעלן ווערן איבריג.
- **3 מעריך: Flan UL2 20B** – א 20 ביליאן פאראמעטער מאדעל איז ארויסגעלאזט געווארן, גאנץ אפן-סאורס און באגרעניצונגען. ער האט א קאנטעקסט פון 2048 טאקענס און א פארבעסערונג ביז 7.4% איבער FlanT5-xxl.
- **6 מעריך: Brave Search Summarizer** – Brave האט לאנסירט א פונקציע וואס פאסט צוזאמען זוך-רעזולטאטן פון פארשידענע קוואלן. זי ניצט פארשידענע שפראך-מאדעלן, אבער נישט OpenAI's GPT, און זי איז צוגעגלייך פאר אלע באניצער.
- **7 מעריך: Google's Universal Speech Model** – דאס איז דער ערשטער רעזולטאט פון Google's איניציאטיוו פאר 1,000 שפראכן. דער מאדעל האט 2 ביליאן פאראמעטער און איז טריינירט געווארן אויף 12 מיליאן שעה רעדע און 28 ביליאן טעקסט-זאצן. ער איבערזעצט יעצט מער ווי 300 שפראכן און ווערט שוין גענוצט ביי YouTube.



2023 מערץ 7 · Universal Speech Model Google

- **7 מערץ: Microsoft Copilot** – Microsoft ניצט AI צו פארמעסטן זיך מיט Oracle, Salesforce און SAP אויפן קארפאראטיוון מארק. דער Copilot ענטפערט קונה-טעלעפאנען, פאסט צוואמען פארקויפונג-מיטינגס און שרייבט מארקעטינג. נאכדעם האט Microsoft לאנסירט Copilot אין Microsoft Graph, וואס פארבינדט שפראך-מאדעלן מיט די באניצער'ס דאטא אין Microsoft Graph.

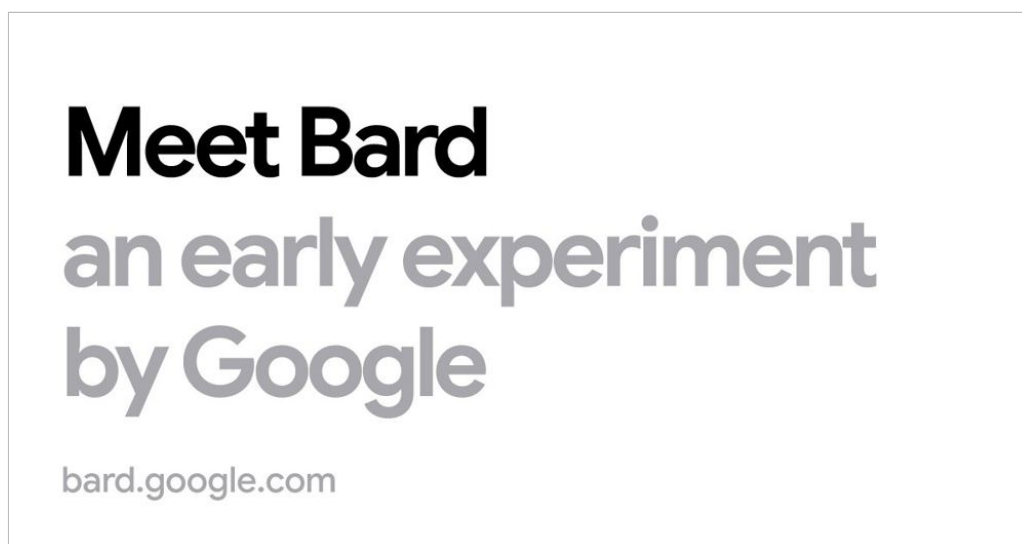


2023 מערץ 7 · Microsoft Copilot

- **9 מערץ: Einstein GPT פון Salesforce** – Salesforce האט לאנסירט Einstein GPT, וואס זי רופט די ערשטע גענעראטיווע AI פאר CRM. עס איז געבויט אין שותפות מיט OpenAI און פארבינדט זיך מיט דאטא פון Salesforce Data Cloud.

- **10 מערץ: Copy.ai און Grammarly** – Copy.ai האט ארויסגעלאזט א טשעט וואס קען אראפנעמען דאטא פון וועבזייטן און צוזאמענפאסן YouTube ווידעאס. Grammarly לייגט צו AI וואס שרייבט אינעם באניצער'ס אייגענעם סטיל.
- **13 מערץ: Claude's Discord** – Discord פרובירט א נייעם Clyde טשעטבאט מיט ChatGPT טעכנאלאגיע. ער ענטפערט פראגעס און פירט שמועסן, און עס קומען אויך AI צוזאמענפאסונגען פון שמועסן.
- **13 מערץ: קלאנגען פאר GPT-4** – Andreas Braun, CTO פון Microsoft Germany, האט געזאגט אז GPT-4 וועט אריינגעפירט ווערן די קומענדיגע וואך, מיט מולטימאדאלע מאדעלן וואס וועלן ערמעגליכן למשל ווידעאס.
- **14 מערץ: OpenChatKit** – Together האט ארויסגעלאזט OpenChatKit, אן אפן-סאורס נאכמאך פון ChatGPT. עס איז א מאדעל מיט 20 ביליאן פאראמעטער, טריינירט אויף 43 מיליאן אינסטרוקציעס. אלע וואגן, קאוד און טריינירונג-דאטא זענען אפן.
- **14 מערץ: Alpaca פון Stanford** – Stanford פארשער האבן ארויסגעלאזט Alpaca, אן אפן מאדעל וואס פאלגט אינסטרוקציעס, כדי די אקאדעמישע וועלט זאל קענען ארבעטן אויף די פראבלעמען פון אזעלכע מאדעלן. ער האט געקאסט ווייניגער ווי \$600 צו מאכן. שפעטער האבן די פארשער אראפגענומען דעם דעמא, צוליב קאסטן און זיכערהייט-זארגן.
- **15 מערץ: Claude's Anthropic** – Anthropic, וואס איז געגרינדעט געווארן פון געוועזענע OpenAI ארבעטער, האט לאנסירט Claude, א קאנקורענט צו ChatGPT. Claude קען זוכן אין דאקומענטן, צוזאמענפאסן, שרייבן, קאוד און ענטפערן פראגעס. Anthropic זאגט אז ער איז "פיל ווייניגער גענייגט צו געבן שעדליכע ענטפערס".
- **15 מערץ: GPT-4** – OpenAI האט ארויסגעלאזט GPT-4, א קליגערער און זיכערער מאדעל, הגם ער האט נאך זיינע באגרעניצונגען. לויט אינערליכע אפשאצונגען איז ער 82% ווייניגער גענייגט צו ענטפערן אויף פארבאטענע בקשות, און 40% מער גענייגט צו געבן פאקטישע ענטפערס ווי Duolingo, GPT-3.5. Khan Academy און Stripe ניצן אים שוין. א שרייבער וואס האט אים אויסגעפרובט האט געפונען אז ער קען א ברייטערן סכום ארבעטן און האט הויך אפגעשניטן אויף טעסטן ווי דער Uniform Bar Exam.
- **15 מערץ: Google PaLM API און Workspace** – Google האט אנאנסירט דעם PaLM API און MakerSuite, וואס ערמעגליכן פראגראמירער צו מאכן פראטאטיפן פון גענעראטיווע אפליקאציעס, ערשט אין א פריוואטע פרובע. Google האט אויך אנאנסירט AI פונקציעס פאר Docs, Gmail, Sheets און Slides.
- **15 מערץ: Wonder Studio** – Wonder Dynamics האט ארויסגעלאזט א געצייג וואס אנימירט, באלייכט און לייגט אריין CG פיגורן אין א לעבעדיגע סצענע, אלעס אינעם בראוזער.
- **16 מערץ: Google Cloud Vertex AI** – Google Cloud האט לאנסירט גענעראטיווע AI אין Vertex AI און דעם Generative AI App Builder פאר ביזנעסער און רעגירונגען. אין צוזאמענהאנג דערמיט האט Google געשלאסן א שותפות מיט Gretel כדי צוצולייגן סינטעטישע דאטא פאר פריוואטקייט.
- **16 מערץ: LinkedIn** – LinkedIn לייגט צו AI געצייג צו שרייבן פראפילן און באשרייבונגען פון דזשאבס.

- **20 מערץ: Midjourney v5** – די נייע ווערסיע שאפט פאטא-רעאליסטישע בילדער. זי ברענגט א ברייטערן סטיל-בארייך, שארפערע רעאקציע צו פראמפטס און צוויי מאל העכערע רעזאלוציע. זי איז צוגעגליך אלס אלפא פאר אבאנענטן.
- **20 מערץ: VALL-E X** – א מאדעל פון Microsoft וואס קלאנירט א קול פון א 4-10 סעקונדע ביישפיל און רעדט דערמיט אין אן אנדער שפראך.
- **21 מערץ: Stability AI – Stable Diffusion Reimagine** – נייער געצייג אין Clipdrop שאפט אן א שיעור ווערסיעס פון איין ארויפגעלייגטן בילד, אן קיין פראמפטס.
- **21 מערץ: Runway Gen-2** – א מולטימאדאלע סיסטעם וואס שאפט ווידעאס פון טעקסט. זי האט עטליכע מאדוסן: טעקסט צו ווידעא, בילד צו ווידעא, סטילזאציע און סטאריבארד.
- **22 מערץ: Google's Bard** – Google האט לאנסירט Bard, א לייכטע ווערסיע פון LaMDA, א שמועס-מאדעל מיט 70 ביליאן פאראמעטער. מען שרייבט זיך איין אויף א ווארט-ליסטע. שפעטער האט Google אפגעליינקט באריכטן אז Bard איז טריינירט געווארן מיט ChatGPT דאטא.



2023 מערץ 22 · Bard's Google

- **22 מערץ: OpenAI שטעלט אפ Codex** – די 4 קאוד-מאדעלן וואס זענען געווען אין א פרייע בעטא זינט 2021 ווערן אפגעשטעלט לטובת turbo און GPT-4. פארשער זארגן אז דאס וועט שטערן איבערצוזורן פריערדיגע פארשונג. OpenAI האט דערנאך געעפנט א פראגראם וואס גיט פארשער סובסידירטן צוטריט און ווייטערדיגן צוטריט צו Codex.
- **22 מערץ: Adobe Firefly** – Adobe האט לאנסירט א פריוואטע בעטא פון Firefly, א משפחה פון אייגענע גענעראטיווע מאדעלן פאר בילדער, ווידעא, דיזיין און 3D. מען וועט קענען צייכענען א טייל פון א בילד און שאפן ווערסיעס, צולייגן סטילן, און רעדאגירן ווידעאס דורך טעקסט.
- **22 מערץ: Bing מיט DALL-E** – אינעם Bing טשעטבאט קען מען יעצט שאפן בילדער.
- **23 מערץ: GitHub Copilot X** – GitHub ברענגט צו Copilot א שמועס-אינטערפעיס, אן אפגרעיד צו GPT-4, פארשלאגן פאר פארריכטן באגס, שרייבן דאקומענטאציע און אויטאמאטישע pull requests.

- **23 מערץ:** Nvidia's פאונדעישן מאדעלן אלס סערוויס - Nvidia קומט אריין אינעם געשעפט פון פאונדעישן מאדעלן פאר פירמעס. די מאדעלן ווערן צוגעפאסט צו אייגענע דאטא, און זיי ארבעטן מיט טעקסט, בילדער, ווידעא און 3D.
- **24 מערץ:** ChatGPT פלאגינס - OpenAI האט אריינגעפירט פלאגינס, וואס לאזן ChatGPT באקומען פרישע אינפארמאציע, רעכענען און ניצן דריטע סערוויסעס. די ערשטע זענען געמאכט געווארן פון Expedia, Instacart, KAYAK, Klarna, OpenTable, Shopify, Slack, Wolfram, Zapier און אנדערע. OpenAI גיט צו אז פלאגינס קענען אויך פארגרעסערן זיכערהייט-ריזיקעס. Stratechery און McCormick באשרייבן ChatGPT ווי א פלאטפארמע וואס קען איינשלינגען אנדערע פראדוקטן.
- **24 מערץ:** Opera - דער Opera בראווער האט צוגעלייגט AI פראמפטס און צוטריט צו ChatGPT און ChatSonic פון א זייט-באר.
- **27 מערץ:** Databricks Dolly - Dolly איז א ביליגער מאדעל וואס פאלגט אינסטרוקציעס ענליך צו ChatGPT. ער איז געבויט אויף אן אפן-סאורס מאדעל פון EleutherAI מיט 6 ביליאן פאראמעטער, און איז צוגעפאסט געווארן מיט אינסטרוקציע-דאטא פון Alpaca.
- **28 מערץ:** Sourcegraph's קאוד-מאדעל - Sourcegraph האט ארויסגעלאזט אן אפענעם קאוד-מאדעל וואס קען ענטפערן פראגעס און שרייבן קאוד לויט א פירמע'ס אייגענעם קאוד. עס איז נישט קלאר וואו די וואגן געפינען זיך.
- **29 מערץ:** OpenFlamingo - Laion האט ארויסגעלאזט א פריימווארק פאר טריינירן Flamingo מאדעלן, און א מאדעל מיט 9 ביליאן פאראמעטער טריינירט אויף 5 מיליאן ביישפילן פון Multimodal C4 און 10 מיליאן פון LAION-2B.
- **30 מערץ:** Cerebras GPT מאדעלן - Cerebras האט ארויסגעלאזט מאדעלן פון 111 מיליאן ביז 13 ביליאן פאראמעטער, טריינירט "קאמפיוט-אפטימאל" אויף איר CS-2 סיסטעם און נישט אויף GPUs. לויט Cerebras האט דאס געדויערט נאר עטליכע וואכן, קעגן 5~ חדשים פאר LLaMA.
- **31 מערץ:** Security Copilot און Zoom IQ - Microsoft האט ארויסגעלאזט Security Copilot אין א פארויסווייז, צו העלפן פארטיידיגער אנטדעקן מוסטערן און רעאגירן שנעלער. Zoom IQ פאסט צוזאמען טשעטס און שרייבט אגענדעס, מיט אייגענע מאדעלן און מיט OpenAI.

פארשונג און טעכניק

- **1 מערץ:** Unum טעקסט-בילד מאדעלן - נייע ViT מאדעלן וואס זענען קאנקורענטפעיג מיט CLIP, בעסער פון אים אין פיל-שפראכיגע פעלער, און פיל שנעלער. די טשעקפוינטס זענען ארויף אויף Hugging Face.
- **1 מערץ:** 3D פון איין בילד - RealFusion בויט א 360° מאדעל פון אן אביעקט פון איין בילד, מיט הילף פון א דיפיוזשאן מאדעל וואס "חלומט אויס" נייע אנבליקן. Zero-1-to-3 טוישט דעם קאמערע-ווינקל פון איין בילד און ווייזט שטארקע אלגעמיינערונג. Make-It-3D שאפט הויך-קוואליטעט 3D פון איין בילד.
- **1 מערץ:** MLLM: "שפראך איז נישט גענוג" - א מולטימאדאלער מאדעל מיט 1.3 ביליאן פאראמעטער, טריינירט אויף 360 ביליאן טאקענס. ער האט באקומען 26% אויף אן IQ טעסט, און 73% zero-shot אויף ImageNet.

- **1 מערץ: ראבאטיק** – CACTI איז א פריימווארק פאר ראבאטן-לערנען אין קיך-סביבות. א ראבאט לערנט זיך ציען קארשן מיט נאר 30 מינוט דאטא. דיפיוזשן אויף באוועגונג-וועגן פארבעסערט שטארק ראבאט-קאנטראל. א RL אגענט לערנט זיך שפילן פיאנא מיט ראבאטישע פינגער.
- **1 מערץ: Scene Diffuser** – א גענעראטיווער מאדעל פאר 3D סצענעס וואס פאראייניגט שאפן, אפטימירן און פלאנירן.
- **2 מערץ: Google גראף-לערנען** – א מעטאד צו איבערטראגן וויסן פון גראפן מיט לעיבלס צו גראפן אן לעיבלס.
- **2 מערץ: קאנטראל אויף בילדער** – Collage Diffusion לאזט קאנטראלירן וואו אביעקטן שטייען און ווי זיי זעען אויס, שיכט ביי שיכט. א צווייטע מעטאד לאזט קאנטראלירן עטליכע געגנטן אויף אמאל, נאך בעסער ווי ControlNet.
- **2 מערץ: RLHF אן מענטשן** – אנשטאט מענטשליכע פידבעק ניצט מען אן אנדער מאדעל צו אפשאצן טאקסישקייט, און דאס דינט אלס באלוינונג. Gretel ניצט א פריוואטקייט-באלוינונג וואס מאכט אז מאדעלן זאלן ווייניגער ארויסלאזן סענסיטיווע דאטא.
- **3 מערץ: Hyena** – Stanford האט פארגעשטעלט אן ערזאץ פאר attention וואס איז 100 מאל שנעלער ביי 64 טויזנט טאקענס. ער גלייכט זיך מיט Transformers ביי שפראך-מאדעלירן, און דער קאוד איז ארויסגעלאזט געווארן.
- **3 מערץ: בילד-רעסטאראציע און StraIT** – א מעטאד צו רעסטאורירן בילדער פון יעדער גרייס מיט דיפיוזשן אן טריינירן. StraIT איז א נישט-אויטארגערעסיווער מאדעל וואס איז בעסער פון דיפיוזשן ביי שאפן בילדער.
- **6 מערץ: מערסט ציטירטע פארשונג-ארבעטן 2022** – Google פירט נאך אלץ, און UC Berkeley פירט צווישן די אוניווערסיטעטן.
- **6 מערץ: טעארעטישע פארשטאנד** – 200 יאר אלטע מאטעמאטיק העלפט ערקלערן ווי נעוראלע נעצן ארבעטן. א פארשונג-ארבעט ערקלערט דיפיוזשן'ס וואג-סכעמעס דורך ELBO.
- **6 מערץ: Grounded decoding** – א שטייגער צו פירן שפראך-מאדעלן אזוי אז די אקציעס זאלן זיין דורכפירבאר פאר א ראבאט. מען פארגלייכט עס מיט SayCan.
- **6 מערץ: Context Clusters און Early Dropout** – CoCs באטראכטן א בילד ווי א זאמלונג פונקטן, און קאנוואלוציע און אן attention. Dropout נאר אינעם אנהייב פון טריינירן פארמינערט underfitting.
- **7 מערץ: פארשידענע גענעראציע** – א דיפיוזשן מאדעל פאר באוועגונג פון מענטשן. StyO פאר סטיליזירן פנימער פון איין בילד. Word-as-Image פאר בוכשטאבן וואס שפיגלען אפ דעם באדייט פונעם ווארט.
- **7 מערץ: ICIL** – לערנען אינסטרוקציעס אינעם קאנטעקסט מיט איין פיקסירטן פראמפט, מיט א פארבעסערונג ביז 9.3%.
- **7 מערץ: Flamingo נאכמאך** – לעקציעס ווי צו באוועלטליגן דייווערדזשענס ביי טריינירן מולטימאדאלע מאדעלן פון 1-2 ביליאן פאראמעטער.
- **8 מערץ: PaLM-E** – Google ווייזט "פאזיטיווע טראנספער" צווישן ראבאט-קאנטראל, ווידעא און טעקסט אין א מאדעל מיט 562 ביליאן פאראמעטער.

- **8 מערץ:** **Prismer – Nerflets** און **Prismer** ניצט באזונדערע "עקספערטן" פאר זעה-שפראך מאדעלן. Nerflets זענען לאקאלע ראדיאנס-פעלדער פאר 3D סצענעס.
- **9 מערץ:** **שפראך-מאדעלן לערנען אנדערש** – קלענערע מאדעלן פארלאזן זיך אויף זייער פאראויסגעוואוסטן וויסן, בשעת גרעסערע לערנען זיך פון נייע ביישפילן. עס איז אויך דא אן אויספירלעכע ערקלערונג וועגן "אויפקומענדיגע פעאיגקייטן".
- **9 מערץ:** **Visual ChatGPT – Microsoft** פארבינדט ChatGPT מיט בילד-מאדעלן צו רעדן, צייכענען און רעדאגירן. דער קאוד איז אפן, אלס א ביישפיל פון ניצן געצייג.
- **9 מערץ:** **X-Avatar און VPD** – א מאדעל פאר פולע דיגיטאלע מענטשן, און א פריימווארק וואס ניצט דיפיוזשאן מאדעלן פאר אויפנעמען בילדער.
- **10 מערץ:** **פראמפטס, ווידעא און סעגמענטאציע** – **Magic Prompt** שאפט אויטאמאטיש פראמפטס. **Video-P2P** רעדאגירט ווידעא. **ODISE** דערגרייכט 23.4 PQ און 30.0 mIoU אויף ADE20K.
- **13 מערץ:** **Masked Image Modeling – PixMIM און DeepMIM** פארבעסערן ביידע MIM טריינירונג.
- **14 מערץ:** **MVImgNet און StyleGANEX** – א דאטאסעט פון 6.5 מיליאן פריימס פון פארשידענע ווינקלען, און א מעטאד צו ניצן StyleGAN אויף נישט-אויסגעגליכענע פנימער.
- **15 מערץ:** **פרי-טריינירן און ראבוסטקייט** – א צוויי-ריכטונגדיגער פרי-טריינירונג פאראדיגמע פאר אויספילן טעקסט. MVP פארבעסערט ראבוסטקייט מיט 8%.
- **16 מערץ:** **UPRISE און Edit-A-Video** – אויטאמאטישער אויסקלייב פון פראמפטס, און רעדאגירן ווידעא דורך טעקסט.
- **17 מערץ:** **S4 – S4, GANformer, NeRF** מאדעלן קענען ווערן א ערזאץ פאר LSTM אין RL. א נאכמאך פון **Re-ReND**, **GANformer** ווייזט NeRF אין רעאלער צייט אויף שוואכע געצייג. **BiFormer** ניצט שפארעוודיגע attention.
- **20 מערץ:** **פראמפטס און Transformers** – אן אויספירלעכער איבערבליק פון פראמפט-אינזשעניעריע. א פארשונג ווייזט אז געוויסע דימענסיעס אינעם residual זענען פיל שטערקער פון אנדערע. **COLT5** פירט אויפן **SCROLLS** בענטשמארק.
- **20 מערץ:** **נאך ארבעטן** – **GlueGen** פאר אפגרעידן טעקסט-צו-בילד מאדעלן. **CoDEPS** פאר ראבאט-נאוויגאציע. **LION** שפארט 11.5% טריינירונג-פאראמעטער. **DiffusionRet** פאר זוכן ווידעא לויט טעקסט.
- **21 מערץ:** **וועלכע דאטא העלפט** – א סימולאטאר וואס שאצט אפ דעם איינפלוס פון יעדע דאטא אויפן מאדעל.
- **22 מערץ:** **SVDiff און CG3D** – **Google**'ס **SVDiff** פאר פערזענליכע בילדער מיט פיל קלענערע פאראמעטער. **CG3D** ברענגט CLIP צו 3D.
- **23 מערץ:** **דעוואליקאציע, FreeDoM, Text to Room** – א פשוטע מעטאד צו אויסמעקן ענליכע דאטא דורך קלאסטערינג. שאפן גאנצע צימערן פון טעקסט. **FreeDoM**, א קאנדיציאנעל דיפיוזשאן מאדעל אן טריינירן.

- **24 מערץ:** "פונקען פון AGI" – Microsoft פארשער טענה'ן אין א גרינטליכער פארשונג-ארבעט אז GPT-4 ווייזט שוואכע פונקען פון אלגעמיינע אינטעליגענץ, נאענט צו מענטשליכן ניוואו.
- **24 מערץ:** **RepoCoder** און **Pix2Video** – אויספילן קאוד אויפן ניוואו פון א גאנצן רעפאזיטאריע, מיט א נייעם בענטשמארק RepoEval. רעדאגירן ווידעא מיט בילד-דיפיוזשאן.
- **27 מערץ:** **רעקאמענדאציעס און דעטעקציע** – א גענעראטיווע רעקאמענדאציע-סיסטעם, מיט ביז 17% פארבעסערונג. א פארשונג-ארבעט ווייזט אז מען קען כמעט נישט דערקענען טעקסט וואס AI האט געשאפן.
- **27 מערץ:** **MagicFusion** און **ReVersion** – צונויפמישן צוויי דיפיוזשאן מאדעלן, און ארויסנעמען א באציאונג צווישן אביעקטן פון בילדער.
- **28 מערץ:** **האלוצינאציעס** – לויט דעם שרייבער איז דער פראבלעם פון האלוצינאציעס אין גרונט א פראבלעם פון אפשאצן אומזיכערקייט.
- **29 מערץ:** **CLIP, Imagen** און **ChatGPT אלס לעיבלער** – EVA-CLIP טריינירט ביליגער. Imagen קאנקורירט מיט CLIP ביים קלאסיפיצירן. gpt-3.5-turbo איז אפט בעסער פון מענטשליכע לעיבלער, און ביליגער מיט בערך 20 מאל.
- **29 מערץ:** **Anti-DreamBooth** און **Text2Video-Zero** – א שוץ דורך ראש וואס שטערט מיסברויך פון פערזענליכע בילדער. א ביליגע מעטאד פאר ווידעא פון טעקסט אן ווידעא-דאטא.
- **30 מערץ:** **בילד, ווידעא, 3D** – StyleDiffusion פאר רעדאגירן בילדער. א ווידעא-דעמא פון 8 פריימס. Instruct 3D-to-3D פון Sony. א מעטאד פאר קאנסיסטענץ אין טעקסט-צו-3D. VIVE3D פאר רעדאגירן פנימער אין ווידעא.
- **31 מערץ:** **פערזענליכע גענעראציע** – DreamBooth3D שאפט 3D פון 3-6 בילדער. HiPer רעדאגירט בילדער אן צופאסן דעם מאדעל. HoloDiffusion טריינירט 3D מאדעלן פון 2D בילדער.

געצייג פאר פראגראמירער און אפן-סאורס

- **1 מערץ:** **Whisper** אויף **iPhone** – Whisper.cpp מאכט OpenAI's Whisper אסאך שנעלער, מיט 4-ביט קוואנטזאציע און א WASM ווערסיע. די קוואנטזאציע ארבעט נאר ביי גרעסערע מאדעלן.
- **1 מערץ:** **פארשידענע געצייג** – T3-Turbo-AI פאר וועב-אפליקאציעס מיט LLMs. Latent Blending פאר אנימאציעס צווישן בילדער. LoRA מאדעלן פאר פאנאראמעס. א זאמלונג פון ChatGPT jailbreaks.
- **2 מערץ:** **PromptLoop** – AI פארמולעס אין Google Sheets און Excel, ווי SUM אדער VLOOKUP.
- **2 מערץ:** **Nvidia** ווידעא-שארפקייט – נייע דרייווערס פאר RTX 30 און 40 קארטלעך שארפן אלטע ווידעאס אין Chrome און Edge.
- **3 מערץ:** **PromptCraft-Robotics** און **Nebullvm** – א קאמיוניטי פאר פראמפטס אין ראבאטיק, און מאדולן צו אפטימירן AI סיסטעמען.
- **6 מערץ:** **טריינירן און ביבליאטעקן** – א טיפ אז gradient accumulation קען שטארק אפשטעלן multi-GPU טריינירן. ביבליאטעקן פאר OpenAI אין Python און PHP. ליסטעס פון AI געצייג.

- **7 מערך:** GPU, OpenICL, LLaMA int8 – לאדן מאדעלן גלייך אויפן GPU. OpenICL פאר לערנען אינעם קאנטעקסט. א קורס פון University of Li ge. LLaMA-13B לויפט אין 24 GiB.
- **8 מערך:** Davinci Functions און מער – געצייג פאר פשוטע אויפגאבן מיט מאדעלן. SGD געשריבן אין Stable Diffusion. SQL אין Photoshop. קאטעגאריע-טעאריע.
- **9 מערך:** Stable Diffusion אינעם בראווער – מיט WebGPU לויפט Stable Diffusion אינגאנצן אינעם בראווער, אן קיין סערווער.
- **9 מערך:** דאטא און ChatGPT אויף Leetcode – א self-instruct דאטאסעט מיט מער ווי 50 טויזנט אינסטרוקציעס. ChatGPT געפרובט אויף אלע Leetcode אויפגאבן. Tsai. Llama-dl.
- **10 מערך:** Kaggle Models – א צענטער פון כמעט 2,000 מאדעלן פון Google, DeepMind און אנדערע, צוזאמען מיט TensorFlow Hub.
- **10 מערך:** LLaMA 4-בייט און Accelerate – GPTQ קוואנטיזאציע שפארט אסאך GPU זכרון. Accelerate שטיצט fp8, PyTorch 2.0 און TPU. א דאטאסעט פון The Office. א ChatGPT wrapper.
- **13 מערך:** LLaMA אויפן לעפטאפ – LLaMA לויפט אויף M1 און M2. מען קען אים אינסטאלירן מיט npx. Text-Generation-Webui. PEFT. RLHF ערלויבט צופאסן ביז 20 ביליאן פאראמעטער אויף א 24GB GPU.
- **14 מערך:** געצייג – Hyper light speed GPT טריינירט א מאדעל אין 3 מינוט. Guardrails פאר סטרוקטור אין ענטפערס. ChatGPTNet. Whisperboard. פאר .NET.
- **15 מערך:** UniDiffuser און Writeout.ai – א פאראייניגטער מולטימאדאלער דיפיוזשאן, און איבערזעצן קול-פייילס מיט Whisper.
- **16 מערך:** PyTorch 2.0 און LLaMA-rs – PyTorch 2.0 איז יעצט סטאביל. LLaMA איז איבערגעטראגן געווארן צו Rust.
- **17 מערך:** NanoT5, LLaMA אין Hugging Face – NanoT5 דערגרייכט ~40 RougeL מיט איין GPU. LLaMA, וועמענס וואגן זענען ארויסגעלעקט, איז יעצט אין OpenAI. Transformers. Alpaca.cpp. Eval's פאר בענטשמארקס.
- **20 מערך:** ChatGPT קאנקורענט פאר \$85 טויזנט – לויט א חשבון קען מען טריינירן א 7 ביליאן מאדעל אין 82 טויזנט A100 שעה. Alpaca Lora פאסט צו LLaMA אויף פשוטע הארדווער. ML באפעלן. Ggml.
- **21 מערך:** Alpaca 30B און מער – LLaMA 30B צוגעפאסט צו פאלגן אינסטרוקציעס. TRL 0.4.1. LangFlow. Bloop זוכט אין קאוד מיט Semantic Kernel. GPT-4.
- **22 מערך:** GPU חשבון און GPTNeoX 2.0 – מען דארף 16 GB × ביליאנען פאראמעטער. GPTNeoX 2.0 איז סינכראניזירט מיט AI Commits. deepspeed. א ליסטע פון אפענע ChatGPT ערזאצן.
- **23 מערך:** ChatLlama און מעדיצין – א טשעט אויף Alpaca 7B. מעדיצינישע גענעראטיווע מאדעלן. Next.js ChatGPT.
- **24 מערך:** ModelScope – א פלאטפארמע פאר פארגלייכן מאדעלן.

- **27 מערך:** NanoPaLM, Shell GPT – NanoPaLM, מיט multi-query attention. Shell GPT פאר דער קאמאנד-ליין.
- **28 מערך:** BlenderGPT און מער – GPT-4 שרייבט קאוד פאר Google, PRESTO, Blender, Anthropic, Pix2Struct, MiniLLM, Cursor פלאגיןס פאר דאטאסעט. פיל-שפראכיגער שמועס-דאטאסעט.
- **29 מערך:** xTuring און Taxy – xTuring פאר צופאסן מאדעלן. Taxy ניצט GPT-4 צו פירן דעם בראווער.
- **30 מערך:** Flux, GPT4All, Lit-LLaMA – Flux ווייזט ענטפערס אין א בויס. GPT4All איז דאטא און קאוד פאר אן אסיסטענט. Lit-LLaMA איז א גאנץ אפענע ווערסיע פון LLaMA.
- **31 מערך:** LLaMA-Adapter – א לייכטע מעטאד צו צופאסן LLaMA מיט 52 טויזנט Alpaca ביישפילן. Kittl AI. AI Library מיט מער ווי 1000 געצייג.

פירמעס, געלט און ביזנעס

- **1 מערך:** Meta און ByteDance – Meta גרינדעט א טיעם פאר גענעראטיווע AI אין אירע אפליקאציעס, טראץ קלאנגען פון אנטלאזונגען. ByteDance בויט GeniusJoy, א לערן-פלאטפארמע.
- **2 מערך:** Typeface – דער סטארטאפ פון דעם געוועזענעם Adobe CTO Abhay Parasnis האט אויפגעהויבן \$65 מיליאן פון Lightspeed, GV, M12 און Menlo.
- **2 מערך:** YouTube און Meta רעקלאמעס – YouTube'ס נײַער פירער Neal Mohan טיזט AI געצייג. Meta'ס AI רעקלאמע ברענגט גוטע רעזולטאטן, אבער מארקעטינג-מענטשן האבן מורא אז זיי פארלירן קאנטראל.
- **3 מערך:** Google Cloud און Ford – דער Google Cloud CEO האט געזאגט פאר די ארבעטער אז "די שפיל ענדיגט זיך נישט אין דער ערשטער מינוט". Ford גרינדעט Latitude AI, מיט 550 ארבעטער אין Pittsburgh פאר זעלבסט-פארנדע קארס.
- **3 מערך:** OpenAI'ס CTO און EleutherAI – א באריכט וועגן דעם וועג פון OpenAI'ס CTO. EleutherAI האט ארויסגעגעבן 28 פארשונג-ארבעטן און 10 מאדעלן.
- **8 מערך:** Salesforce פאנד – א פאנד פון \$250 מיליאן. די ערשטע אינוועסטירונגען זענען אין Cohere, Anthropic, You.com און HealthAI.
- **8 מערך:** Stability קויפט ClipDrop – ClipDrop האט מער ווי 15 מיליאן באניצער. אין קאמפעיטיטיוון ML ניצן 96% פון די לייזונגען PyTorch.
- **8 מערך:** Scale AI – א פירמע פאר טריינינג-דאטא, געגרינדעט פון Alexandr Wang ווען ער איז געווען 19 יאר אלט. זי ווערט יעצט אפגעשאצט מער ווי \$7 ביליאן.
- **9 מערך:** ווי Google איז געווארן פארזיכטיג – Google פארשער האבן געדרוקט ארויסצולאזן א טשעטבאט מיט צוויי יאר צוריק, אבער די פירמע האט זיך געהאלטן פארזיכטיג און דערמיט געגעבן Microsoft א געלעגנהייט.
- **13 מערך:** ChatGPT אין Azure – Microsoft האט אריינגעשטעלט דעם ChatGPT API אין Azure OpenAI Service.

- **14 מערץ:** GM – GM פלאנט א ChatGPT אסיסטענט אין אירע קארס, דאנק איר שותפות מיט Microsoft.
- **15 מערץ:** Morgan Stanley – די באנק ניצט GPT-4 פאר אן אינערליכן טשעטבאט וואס געפינט אינפארמאציע פאר אירע עצה-געבער.
- **16 מערץ:** Adept – Adept האט אויפגעהויבן \$350 מיליאן אין Series B, אנגעפירט פון General Catalyst און Spark Capital, כדי צו אויטאמאטיזירן ווייכווארג-פראצעסן.
- **17 מערץ:** Stripe און PwC – Stripe העלפט OpenAI מאכן געלט פון אירע פראדוקטן, און ניצט GPT-4. PwC האט א גלאבאלע שותפות מיט Harvey, א לעגאל AI וואס ווערט געשטיצט פונעם OpenAI Startup Fund.
- **20 מערץ:** Zapier, Khan, Duolingo – Zapier's NLA API גיט צוטריט צו 5,000 אפליקאציעס און 20,000 אקציעס דורך נאטירלעכער שפראך. Khan Academy's Khanmigo ארבעט מיט GPT-4. Duolingo Max ברענגט Role Play און Explain my Answer.
- **20 מערץ:** AssemblyAI – דער נייער Conformer ASR מאדעל איז צוגענגליך נאר דורך API.
- **22 מערץ:** Nvidia H טשיפס – נייע טשיפס מיט 188 GB זכרון און 990 TFLOPS.
- **23 מערץ:** Mozilla.ai – Mozilla גרינדעט Mozilla.ai פאר אפענע, "געטרייע" AI. אויך האט Mozilla געעפנט א Responsible AI טשעלענדזש.
- **23 מערץ:** Ubisoft – Ubisoft ניצט Ghostwriter צו שרייבן דיאלאג פאר שפילן.
- **27 מערץ:** Character AI – א נייער מאדעל און א נייע געלט-רונדע.
- **28 מערץ:** Nvidia's Jensen Huang – ער רופט ChatGPT דעם "iPhone מאמענט" פון AI, און זאגט אז דאס אינפערענס-געשעפט האט געמאכט א שפּרונג.
- **28 מערץ:** Apple קויפט WaveOne – א סטארטאפ פאר ווידעא-קאמפרעסיע מיט AI.
- **29 מערץ:** Perplexity און Replit – Perplexity האט לאנסירט אן iPhone אפליקאציע, האט אויפגעהויבן א Series A, און האט 2 מיליאן באניצער. Google טוט זיך צוזאמען מיט Replit's Ghostwriter קעגן GitHub Copilot.
- **29 מערץ:** Roblox – Roblox ניצט גענעראטיווע AI פאר שאפן שפילן.
- **30 מערץ:** Microsoft באגרעניצט Bing – Microsoft באגרעניצט דעם צוטריט צו Bing פאר קאנקורירנדע טשעטבאטס.
- **31 מערץ:** Midjourney – Midjourney שטעלט אפ די פרייע פרובעס, נאך א פלוס פון נייע באניצער.

רעגירונגען, געזעצן און זיכערהייט

- **2 מערץ:** אומדירעקטע פראמפט-אינדזשעקציע – א שוואכקייט אין Bing Chat: אן אנגריפער קען באהאלטן אינסטרוקציעס אויף א וועבזייט וואס דער באניצער האט אפן, און Bing ווערט דערמיט אן אינסטרומענט צו ארויסלאקן פערזענליכע אינפארמאציע.
- **3 מערץ:** רומעניע'ס "Ion" – א רעגירונג-פלאטפארמע וואס נעמט אויף די מיינונגען פון בירגער כדי צו פירן פאליס.

- **7 מערץ:** שווינדל מיט AI קולות – אין 2022 האבן שווינדלער גע'גנב'עט \$11 מיליאן דורך זיך מאכן פאר קרובים מיט AI קולות.
- **15 מערץ:** Microsoft **אנטלאזט עטיק-טיעם** – אינערהאלב פון 10,000 אנטלאזונגען האט Microsoft אנטלאזט איר גאנצן עטיק און געזעלשאפט טיעם פון AI.
- **28 מערץ:** FTC – די FTC גיט אכטונג אז גרויסע טעכנאלאגיע-פירמעס זאלן נישט באהערשן דעם פעלד.
- **29 מערץ:** **סענסיטיווע דאטא** – ארבעטער לייגן אריין קארפאראטיווע דאטא אין ChatGPT.
- **30 מערץ:** **בריוו פאר א 6-חדשים'דיגע פויזע** – Elon Musk, Steve Wozniak און אנדערע פאדערן א פויזע פון 6 חדשים אין AI. דער בריוו איז ארגאניזירט פון Future of Life Institute, און ער ווארנט פון טיפע ריזיקעס פאר דער געזעלשאפט.

אנאליזן און מיינונגען

- **2 מערץ:** **בויען וועלטן מיט GPT** – א שרייבער ניצט GPT צו שאפן א שטאט, פון בנינים ביז מענטשן.
- **6 מערץ:** **ויי ChatGPT איז געבויט געווארן** – שמועסן מיט פיר פון זיינע בויער, צווישן זיי John Schulman און Jan Leike.
- **6 מערץ:** **AI און ארבעטלאזיגקייט** – Marc Andreessen טענה'ט אז AI וועט נישט ברענגען מאסן-ארבעטלאזיגקייט. אן אנדער ארטיקל טענה'ט אז דער מערב ליידט פון צו ווייניג אויטאמאציע.
- **7 מערץ:** **Waluigi עפעקט** – נאכדעם וואס מען טריינירט א מאדעל אויף א געוויסער אייגנשאפט, ווערט עס גרינגער ארויסצוברענגען פון אים די פארקערטע אייגנשאפט. אויך איז דא א באריכט וועגן דעם AI Objective Institute.
- **8 מערץ:** **סטארטאפס און וועניטשור קאפיטאליסטן** – ווי AI סטארטאפס זאלן טראכטן וועגן שוץ פון קאנקורענץ. וועניטשור קאפיטאליסטן זארגן אז AI זאל נישט איבערחזר'ן קריפטא.
- **9 מערץ:** **פראנטענד** – ווי AI טוישט פראנטענד פראגראמירן.
- **10 מערץ:** **Anthropic וועגן זיכערהייט** – Anthropic טיילט איר פארשונג אין פעאייגקייטן, צופאסן פעאייגקייטן, און alignment.
- **10 מערץ:** **וועב ווי א דאטאבאזע** – AI קען דערמעגליכן דעם אלטן חלום פון Semantic Web אגענטן.
- **13 מערץ:** **Stable Diffusion LLMs מאמענט** – Simon Willison באשרייבט ווי שפראך-מאדעלן ערלעבן א כוואליע פון חידושים, ענליך צו Stable Diffusion.
- **14 מערץ:** **"Hot mess" טעאריע** – א טעאריע וועגן ריזיקעס פון AI alignment.
- **15 מערץ:** **Gary Marcus** – ער ווארנט פון אן אומבאגרייפער כמות פאלשע אינפארמאציע, און טענה'ט אז היינטיגע מאדעלן האבן שוין ערנסטע ריזיקעס, אפילו אן AGI.
- **16 מערץ:** **Siri, Alexa, און ווי צו אפשטעלן AI** – וויאזוי די קול-אסיסטענטן האבן פארלוירן די AI ראסע. א שרייבער טראכט אויס וועגן צו פארלאנגזאמען AI.

- **20 מערך:** Sam Altman – Altman ווארנט פון פאלשע אינפארמאציע און סייבער-אנגריפן, און זאגט אז עס איז וויכטיג צו פרובירן מיט ריכטיגע באניצער בשעת די סכנות זענען נאך קליין. ער רעדט אויך וועגן ביאס אין דער דאטא, און אין א פאדקאסט מיט Lex Fridman.
- **21 מערך:** ארבעט-מארק – א פארשונג-ארבעט שאצט אפ אז פיזישע און שעפערישע דזשאבס זענען ווייניגער באדראט. Goldman Sachs שאצט אז 300 מיליאן דזשאבס קענען באוירקט ווערן.
- **22 מערך:** פארלאנגזאמען און רייזע – Vox טענה'ט פאר פארלאנגזאמען AI. a16z זעט געלעגנהייטן אין רייזע.
- **23 מערך:** פאלשע אינפארמאציע און מאנאפאל – Bing האט געזאגט אז Bard איז פארמאכט געווארן, באזירט אויף פאלשע קוואלן. Politico פרעגט צי AI וועט ווערן א מאנאפאל.
- **24 מערך:** Bill Gates – Gates רופט AI די וויכטיגסטע טעכנאלאגישע אנטוויקלונג זינט דעם גראפישן אינטערפעס, מיט פאטענציאל פאר געזונט, חינוך און קלימאט. Modular טענה'ט אז צעשפליטערטע הארדווער האלט אפ דעם פארשריט.
- **27 מערך:** לאקאלע מאדעלן און Forerunner – א פערזענליכער אסיסטענט אויף אייגענער הארדווער ווערט נעענטער. Brian O'Malley פון Forerunner פארגלייכט דעם היינטיגן מאמענט מיט Web3.
- **28 מערך:** GPT-4 שרייבט א בוך – GPT-4 האט געשריבן א בוך פון 115 זייטן דורך א סעריע פראמפטס.
- **31 מערך:** נאך אנאליזן – Stratechery וועגן "צופעליגע" קאנסומער פירמעס. Sequoia וועגן געצייג פאר פראגראמירער. Jailbreakers טענה'ן אז זיי פארהיטן א פארמאכטע AI צוקונפט. א פארגלייך פון דער היינטיגער AI תקופה מיטן קאמבריאן פעריאד.

ווייטערע ידיעות

- **7 מערך:** ברוסט-קענסער – AI דערקענט קענסער כאטש אזוי גוט ווי ראדיאלאגן. אונגארן איז געווארן א הויפט-פלאץ פאר אויספרובירן.
- **13 מערך:** COVID וואקסין – א וואקסין קעגן אלע וואריאנטן איז אנטוויקלט געווארן מיט הילף פון ML.
- **14 מערך:** עסן-קריזיס – AI קען פאראויסזאגן אויסברוכן פון הונגער פיל פריער.
- **15 מערך:** Cornell קורס – אן אפענער קורס פון Applied Machine Learning.
- **17 מערך:** Vesuvius טשעלענדזש – \$250,000 פרייז, צווישן זיי א \$150,000 הויפט-פרייז, פאר ליינען אן אלטע Herculaneum מגילה מיט ML.
- **17 מערך:** פראטאינען – AI Meta'ס זאגט פאראויס די סטרוקטור פון הונדערטער מיליאנען פראטאינען. DeepMind און Meta ארבעטן מיט ביאלאגישע דאטא פאר מעדיצינען.
- **24 מערך:** קלאנירן זיך – ווי גרינג עס איז צו קלאנירן זיך מיט גענעראטיווע AI.
- **27 מערך:** זוכן לעבן אויף Mars – AI קען העלפן אנאליזירן דאטא פון Mars אין זוכן פון לעבן.

אפריל 2023 איז געווען א חודש פון א פלייץ אפן-סאורס מאדעלן און דאטא: Dolly 2 פון Databricks, Vicuna, Koala, StableLM פון HuggingFace Chat, RedPajama, Stability AI, און O'Meta, Segment Anything, וועלכער האט ארויסגערופן א גאנצע קוואליע פון געצייג. AI אנענטן ווי Auto-GPT און BabyAGI זענען געווארן דער נייסטער טרענד. רעגירונגען האבן אנגעהויבן אריינמישן: איטאליע האט פארבאטן ChatGPT, כינע האט ארויסגעגעבן פלאן-רעגולאציעס, און אייראפע און אמעריקע קלערן געזעצן. Google האט צוזאמענגעשטאמלט DeepMind מיט Brain און לאנסירט די PaLM API. Amazon האט ארויסגעלאזט Elon Musk. Bedrock האט גענומען בויען זיין אייגענע AI, און וועקטאר-דאטאבאזעס ווי Chroma און Pinecone האבן אויפגענומען גרויסע געלט-רונדעס.

מאדעלן און נייע לאנסירונגען

- **3 אפריל: Microsoft Security Copilot** – Microsoft האט פארגעשטעלט Security Copilot, א AI געצייג פאר זיכערהייט-אנאליסטן. עס העלפט שנעל רעאגירן אויף סכנות, פראצעסירן סיגנאלן מיט מאשין-שנעלקייט און אפשאצן ריזיקעס אין מינוטן. דערווייל איז עס אין א פאר-ווייזונג און נאך נישט ארויסגעלאזט פארן פובליק.
- **3 אפריל: Bard ווערט פארבעסערט** – Sundar Pichai האט געזאגט אז Bard וועט אריבערגיין פון א LaMDA-באזירטן מאדעל צו דעם גרעסערן PaLM. שפעטער אין חודש האט Bard אנגעהויבן העלפן שרייבן קאוד אין א ברייטע צאל שפראכן, כאטש ביים לאנסירן האט ער אפגעזאגט ענטפערן אויף קאוד-פראגעס. עס איז נאך נישטא קיין אפיציעלע בענטשמארק-באשטעטיגונג ווי ער שטייט אקעגן אנדערע מאדעלן.
- **5 אפריל: PaLM API און MakerSuite** – Google האט אנאנסירט די PaLM API און MakerSuite, גרינגע געצייג פאר פראגראמירער צו פראבירן און בויען אפליקאציעס מיט גענעראטיווע AI. אין אנפאנג וועט עס זיין צוגענגליך נאר פאר אויסגעקליבענע פראגראמירער, דורך א "Private Preview".
- **6 אפריל: Segment Anything פון Meta** – Meta האט ארויסגעלאזט Segment Anything, א פולשטענדיג אפן-סאורס מאדעל וואס אפטיילט אביעקטן אין בילדער. מען האלט עס פאר א גרויסן שפונג פארויס: Meta האט צונויפגעזאמלט א ריזיגע דאטא-זאמלונג און באקוועמערע געצייג צו צייכענען די דאטא. דער מאדעל איז ליכטיג גענוג צו לויפן גלייך אינעם בלעטערער. דערפון זענען שוין געבוירן געווארן עטליכע געצייג: Chrome-צוגאבן (Magic Copy) וואס קאפירן אן אביעקט פון א בילד גלייך צום קלעפבלאט; קאמבינאציעס מיט Stable Diffusion, Dino, און Blip צו דערקענען, צופאסן און באשרייבן בילדער; Track Anything צו נאכפאלגן אביעקטן אין ווידיאו; SA3D פאר 3D אפטיילונג אין בערך 2 מינוט; און EditAnything און Inpaint Anything, וואס לאזן אראפנעמען, אנפילן אדער אויסטוישן טיילן פון א בילד.

IDEA-Research/
Grounded-Segment-...

idea

INTERNATIONAL
DIGITAL
ACADEMY
粤港澳大湾区数字经济研究院

Grounded SAM: Marrying Grounding DINO with
Segment Anything & Stable Diffusion & Recognize
Anything - Automatically Detect , Segment and...

58

Contributors

304

Issues

18

Discussions

18k

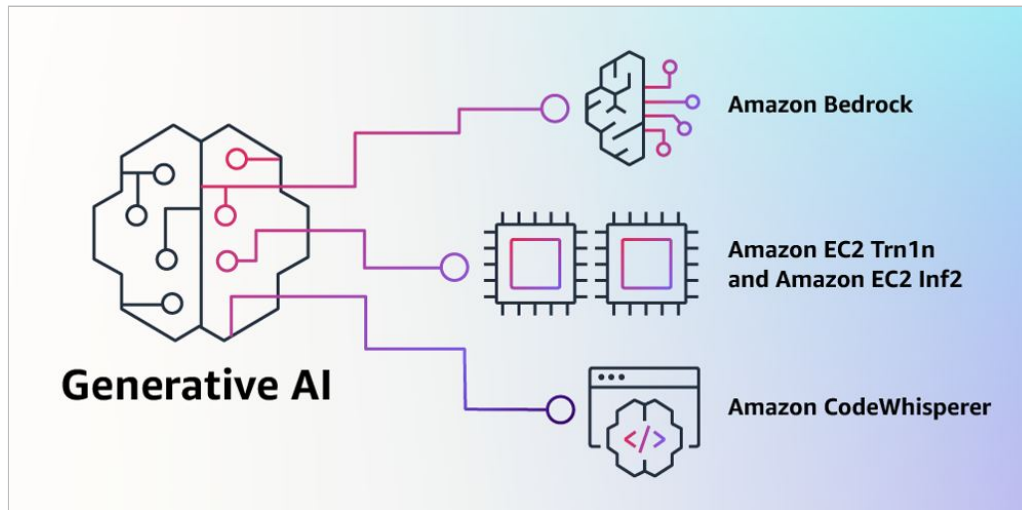
Stars

2k

Forks

2023 אפריל 6 · Meta פון Segment Anything

- 11 אפריל: Poe פון Quora – Quora** האט ארויסגעלאזט Poe, אן אפליקאציע וואס גיט צוטריט צו פארשידענע AI טשעטבאטס אויף איין ארט.
- 11 אפריל: ChatSpot 3 – ChatSpot Alpha 3** ניצט יעצט GPT-4 און קען באהאנדלען ביז 32,000 טאקענס. עס האט א ביבליאטעק פון מוסטער-פראמפטס, קען צוזאמענפאסן ארטיקלען, שטיצט HubSpot פראמפטס, און פירט טיפערע פארשונג איבער פירמעס.
- 11 אפריל: Image Creator אין Edge – Microsoft** עפנט פאר אלע Image Creator אינעם Edge בלעטערער. דאס געצייג שאפט בילדער פון א טעקסט-באשרייבונג מיט OpenAI's DALL-E.
- 12 אפריל: Alibaba's טשעטבאט – Alibaba** פלאנירט אריינצושטעלן איר ChatGPT-קאנקורענט אין אלע אירע פראדוקטן.
- 13 אפריל: Dolly 2 פון Databricks – Databricks** צוויי וואכן נאך Dolly 1.0 האט Databricks ארויסגעלאזט Dolly 2. דאס איז אן אפענער מאדעל, געבויט אויף EleutherAI's Pythia, און טריינירט אויף א נייער דאטא-זאמלונג וואס מען מעג ניצן פאר ביונעס. עס איז נישט קלאר ווי ער פארגלייכט זיך מיט Llama- מאדעלן אדער Ars Technica GPT-3.5 turbo רופט עס א "גאר גרויסע זאך" ווייל עס איז פריי און אפן-סאורס.
- 14 אפריל: AWS Bedrock – Amazon** האט אנאנסירט Bedrock, א סערוויס אויף AWS וואס גיט צוטריט צו יסוד-מאדעלן פון פירמעס ווי Anthropic און Stability AI. מען קען צופירן די אייגענע דאטא פון S3, און אפשר לויפן די מאדעלן אין דער אייגענער AWS אינפראסטרוקטור. דאס איז אטראקטיוו לויט די היינטיגע זארגן איבער דאטא-פרייוואטקייט.



2023 אפריל 14 · AWS Bedrock

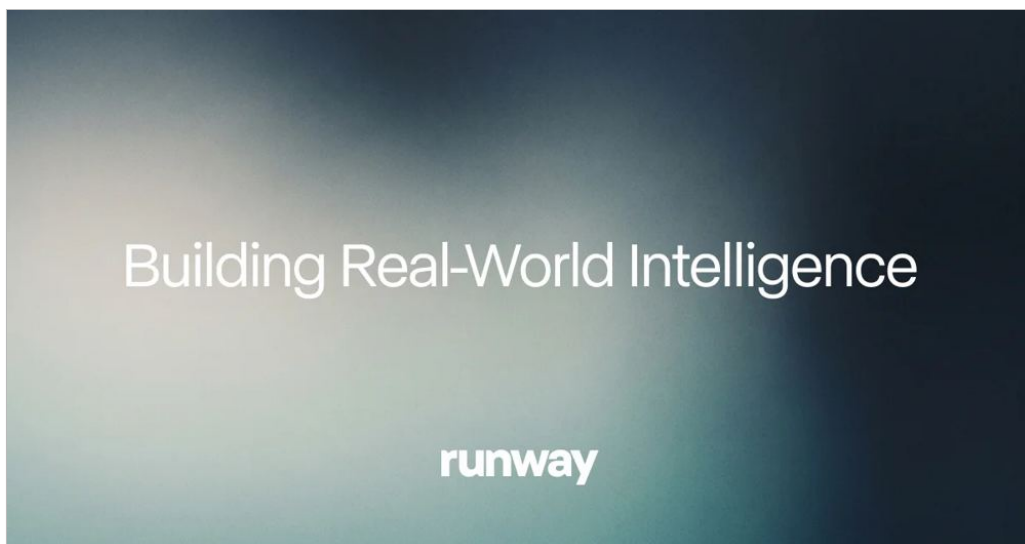
- **14 אפריל:** Vicuna-13B – Vicuna-13B איז אן אפן-סאורס טשעטבאט, טריינירט דורך פיינ-צופאסן LLaMA אויף געשפרעכען וואס באניצער האבן געטיילט. ער שלאגט LLaMA און Alpaca אין איבער 90% פון פעלער, און דערגרייכט איבער 90% פון דער קוואליטעט פון ChatGPT און Bard. דאס טריינירן האט געקאסט בערך \$300. דער קאוד איז אפן פאר נישט-קאמערציעלע באנוץ.
- **17 אפריל:** Paella פון Laion – Laion האט ארויסגעלאזט Paella, א פשוט'ער טעקסט-צו-בילד מאדעל וואס יעדער זאל קענען פארשטיין. זיין סעמפלינג-קאוד איז נאר 12 שורות לאנג.
- **18 אפריל:** פיל-שפראכיגער בילד-מאדעל – א נייער טעקסט-צו-בילד מאדעל ארבעט מיט עטליכע שפראכן, טריינירט אויף LAION HighRes און אן אייגענער דאטא-זאמלונג. ער האט מער פאראמעטערס ווי Stable Diffusion.
- **18 אפריל:** Adobe Firefly און Adobe Lightroom – Adobe ארבעט אויף Firefly AI פארבעסערונגען פאר אירע ווידיאו- און אודיא-אפליקאציעס אין Creative Cloud. אין Lightroom האט זי צוגעלייגט AI פונקציעס ווי Denoise, Curves אין מאסקירונג, אדאפטיווע פריסעטס, און אויסקלייבן מענטשן.
- **20 אפריל:** StableLM פון Stability AI – די שאפער פון Stable Diffusion האבן ארויסגעלאזט StableLM שפראך-מאדעלן מיט 3B און 7B פאראמעטערס און א קאנטעקסט פון 4k טאקענס. ביז יעצט זענען זיי טריינירט אויף בערך 800B טאקענס, מיטן ציל צו דערגרייכן 1.5T. זיי זענען ערלויבט פאר קאמערציעלע באנוץ.
- **20 אפריל:** ווידיאו פון Nvidia – Nvidia'ס פארשער ווענדן "לאטענטע" דיפוזיע, וואס האט געהאלפן ביי בילדער, אויף ווידיאו. מיט קליינע ענדערונגען, וואס בינדן די וועקטארן צוזאמען איבער צייט, באקומט מען זייער קאנסיסטענטע סצענעס.
- **21 אפריל:** RedPajama – Together האט ארויסגעגעבן א דאטא-זאמלונג פון 1.2 טריליאן טאקענס, וואס נאכמאכט די מישונג פון Llama'ס טריינירן. ציל איז צו פארמיידן Llama'ס ליצענץ-באגרענעצונגען. 72 שעה שפעטער האט Mosaic טריינירט אויף דעם דעם ערשטן RedPajama מאדעל: 1B פאראמעטערס, אויף 440 GPUs, איבער 200B טאקענס, מיט AliBi.



The AI community building the future

2023 אפריל 21 · RedPajama

- **21 אפריל:** Atlassian Intelligence – Atlassian האט פארגעשטעלט Atlassian Intelligence, א "חבר אין דער גרופע" פאר אירע קלאוד-פראדוקטן. עס ניצט OpenAI'ס מאדעלן צו שאפן און צוזאמענפאסן אינהאלט, ענטפערן פראגעס, און לערנען זיך די פירמע'ס טערמינען.
- **21 אפריל:** Bark פון Suno – Bark איז א טעקסט-צו-אודיא מאדעל. ער שאפט רעאליסטישע רייד אין פארשידענע שפראכן, ווי אויך מוזיק, הינטערגרונט-ראש און סאונד עפעקטן.
- **25 אפריל:** Gen-1 אויף iOS – Runway האט ארויסגעלאזט אן iOS אפליקאציע מיט Gen-1, וואס פארוואנדלט ווידעאס פון טעלעפאן-קאמערע.



2023 אפריל 25 · Gen-1 אויף iOS

- **25 אפריל:** Sec-PaLM – Google האט ארויסגעלאזט Cloud Security AI Workbench, באזירט אויף Sec-PaLM, א שפראך-מאדעל ספעציעל פאר סייבער-זיכערהייט.
- **26 אפריל:** HuggingFace Chat – HuggingFace האט ארויסגעלאזט א טשעט-אפליקאציע ענלעך צו ChatGPT, באזירט אויף די 30B OpenAssistant Llama מאדעלן. זי איז לאנגזאמער ווי די קאנקורענטן, אבער שטארק פאר א כמעט-אפענעם מאדעל.

- **26 אפריל: ChatGPT פריוואטקייט** – באניצער פון ChatGPT קענען יעצט אויסלעשן די היסטאריע פון זייערע געשפרעכן, אזוי אז מען זאל זיי נישט ניצן צו טריינירן. OpenAI ארבעט אויך אויף ChatGPT Business פאר פירמעס וואס דארפן מער קאנטראל איבער זייער דאטא.
- **26 אפריל: MiniGPT-4 און LLaVA** – LLaVA פאראייניגט א בילד-ענקאדער מיט א שפראך-מאדעל, טריינירט אויף אנווייזונגען וואס GPT-4 האט געשאפן. אין צוזאמענארבעט מיט GPT-4 האט עס דערגרייכט 92.53% אויף Science QA. MiniGPT-4 אויך דאס פארשטיין פון בילד און שפראך צוזאמען.
- **27 אפריל: Upscaling פון Stability AI** – Stability AI האט ארויסגעלאזט אן Upscaling API מיט צוויי אפגעמאדעלטן, SD Latent Upscaler און ESRRGAN, וואס פארגרעסערן בילדער פון 512x512. עס איז אריינגעשטעלט אין די Photoshop און Blender צוגאבן.
- **27 אפריל: DeepFloyd IF** – Stability AI האט ארויסגעלאזט דעם קאוד פון DeepFloyd IF, א מאדעל מיט עטליכע שטאפלען און א פארפרוירענעם T5 טעקסט-ענקאדער. ער קען שרייבן טעקסט אין בילדער אין זייער הויכער קוואליטעט. אין אנפאנג וועט ער ווארשיינליך זיין נאר פאר פארשונג.

פארשונג און טעכניק

- **3 אפריל: GPTEVAL** – א נייע מעטאד ניצט GPT-4, מיט "chain-of-thought" און אויספילן פארמען, צו אפשאצן די קוואליטעט פון געשאפענעם טעקסט. די רעזולטאטן שטימען בעסער מיט מענטשליכע מיינונגען, און די ארבעט ווייזט אויך אז אזעלכע אפשאצער קענען זיין פארשטעלט.
- **3 אפריל: לענגערע קאנטעקסט** – Stanford ארבעט צו פארגרעסערן שטארק דעם קאנטעקסט פון שפראך-מאדעלן. דאס קען עפענען טירן צו צוזאמענפאסן גאנצע ביכער, ענדערן קאוד אין א גאנצן רעפאזיטאריום, און שאפן ווידאו. א צווייטע ארבעט האט געברענגט אן RNN צו א מיליאן טאקענס קאנטעקסט.
- **3 אפריל: בילד-דיפוזיע** – Diff- ∞ איז א דיפוזיע-מאדעל וואס ארבעט מיט אומבאגרעניצטער רעזאלוציע און קען שאפן בילדער שארפער ווי די טריינינג-דאטא. אן אנדערע ארבעט שאפט 3D-באוואוסטזיניגע בילדער פון פארשידענע ווינקלען מיט 2D דיפוזיע.
- **4 אפריל: מאדעלן קריטיקירן זיך אליין** – גרויסע שפראך-מאדעלן קענען קריטיקירן זייער אייגענעם ארויסגאנג און אזוי שטארק פארגעסערן די קוואליטעט. פארשער האבן געפונען אז ווען מען בעט GPT-4 צו קריטיקירן זיין ארבעט, שטייגט די לייסטונג מיט 30%.
- **4 אפריל: ChatGPT ווי דיריגענט** – א פארשונג-ארבעט ניצט ChatGPT ווי א קאנטראלירער, און יעדן Hugging Face מאדעל ווי א ספעציעלן געצייג. דאס איז אן ערשטער שריט צו אויטאמאטיזירן אויפגאבן מיט עטליכע שריט.
- **4 אפריל: VideoMAE V2 און InceptionNeXt** – VideoMAE V2 בויט ווידאו יסוד-מאדעלן מיט "דאפלטער מאסקירונג" און דערגרייכט די בעסטע רעזולטאטן אויף עטליכע דאטא-זאמלונגען. InceptionNeXt צעטיילט גרויסע קאנוואלוציעס אויף פיר צווייגן, פאר שנעלערן טריינירן און בעסערע פינקטלעכקייט.

- **5 אפריל: ראבאטן** – Meta האט טריינירט איין אפן-סאורס Visual Cortex מאדעל וואס באהאנדלט 17 שווערע אויפגאבן, א שריט צו פארבעסערן ראבאטן וואס באוועגן און פירן זאכן.
- **5 אפריל: RLHF** – EleutherAI ערפארשט ווי RLHF ענדערט דאס אינעווייניגסטע פון א קליינעם GPT מאדעל. Hugging Face ווייזט דאס פראקטישע RLHF-טריינירן מיט StackLLaMA. Microsoft DeepSpeed Chat פירט אדורך אלע דריי שטאפלען פון RLHF, און RRHF איז א פשוט'ערע אלטערנאטיוו צו PPO מיט ענלעכע רעזולטאטן.
- **5 אפריל: Koala פון Berkeley** – Koala איז טריינירט אויף דיאלאג- און אנווייזונג-דאטא פון פרייע מקורות אויף דער וועב, נישט פון OpenAI מאדעלן. ער ווייזט זיך ענלעך צו שטארקע פארמאכטע מאדעלן, און דער קאוד מיט דער דאטא זענען אפן.
- **5 אפריל: ווידיאו פארשונג** – Follow Your Pose שאפט ווידיאוס פון פיגורן מיט קאנטראלירטע פאזעס. ReBotNet פארבעסערט לייוו ווידיאו מיט 50 בילדער א סעקונדע און בערך 5GB זכרון, אויפן שטאפל פון מאדעלן 8 מאל גרעסער. VideoCrafter איז אן אפן-סאורס געצייג-קאסטן פאר טעקסט-צו-ווידיאו. DreamPose, VidStyleODE, AdaM און Soundini באהאנדלען בילד-צו-ווידיאו, רעדאקטירן ווידיאו, אפטיילן מענטשן פונעם הינטערגרונט, און עפעקטן געפירט דורך קלאנג.
- **6 אפריל: LLM איבערבליק און Pythia** – א ברייטער איבערבליק באשרייבט די אנטוויקלונג פון גרויסע שפראך-מאדעלן. EleutherAI האט ארויסגעלאזט הונדערטער שטאפלען (checkpoints) פון Pythia מאדעלן, אזוי אז מען קען שטודירן וואס א מאדעל לערנט זיך במשך פונעם טריינירן.
- **6 אפריל: אודיא** – AUDIT רעדאקטירט קלאנג לויט אנווייזונגען און דערגרייכט די בעסטע רעזולטאטן. CLAP, א מאדעל ענלעך צו CLIP אבער פאר אודיא, האט ארויסגעלאזט טריינירטע ווייטס.
- **7 אפריל: ראייה (Vision)** – א מאדעל ניצט טאקענס פון געמישטע רעזאלוציע פאר Vision Transformers. RECLIP טריינירט CLIP מיט 6-8 מאל ווייניגער רעכן-קראפט. Meta's DINOv2 טריינירט א ViT מאדעל מיט 1 ביליאן פאראמעטערס אן השגחה, און די קלענערע מאדעלן וואס שטאמען דערפון שלאגן קיימדיגע מעטאדן. SpectFormer, RT-DETR (ביז 54.8% AP), שלאגט YOLO אין רעאל-צייט און SparseFormer ברענגען נייע ארכיטעקטורן.
- **10 אפריל: טריינירן גרויסע מאדעלן** – W&B האט א באריכט מיט עצות פאר טריינירן מאדעלן פון איבער 1B פאראמעטערס. EleutherAI's "Transformer Math" רעכנט אויס וויפיל GPUs מען דארף. איין מסקנא: מען זאל נישט טריינירן אויף ווייניגער ווי 200B טאקענס, אבהענגיג נישט פון דער גרייס.
- **10 אפריל: באשטעטיגן קאוד** – א פארשונג-ארבעט שאפט מאטעמאטישע באווייזן וואס א פארמעלע סיסטעם קען קאנטראלירן, און פאררעכטן זיי. אן אנדערע מעטאד שאפט קודם טעסטן און דערנאך קאוד, און האט אזוי דערגרייכט א רעקארד אויף HumanEval, שוין איבערגעיאגט דורך Reflexion.
- **10 אפריל: רעדאקטירן בילדער** – Inst-Inpaint נעמט אראפ אביעקטן פון בילדער לויט ווערטער, אן מאסקעס. C-LoRA פארמיידט "פארגעסן" ביים צולערנען נייע באגריפן. Collaborative Diffusion שאפט פנים'ער מיט עטליכע סארטן באדינגונגען. DiffFit פאסט צו דיפוזיע-מאדעלן ביליג, און האט באקומען א רעקארד FID אויף ImageNet 512x512.
- **10 אפריל: Gaussian processes** – Google פאר-טריינירט Gaussian process מאדעלן, אזוי אז באיעסיאנישע אפטימיזאציע אויף נייע פראבלעמען זאל גיין שנעלער.

- **11 אפריל:** גענעראטיווע אגענטן – פארשער האבן געשטעלט שפראך-מאדעלן אנצופירן א גאנצע שטאט פון סימולירטע מענטשן. די אגענטן האבן ארויסגעוויזן מענטשן-ענלעכע סאציאלע אויפפירונג: זיי האלטן א סדר-היום, פלאנירן אקטיוויטעטן און דערמאנען זיך אין פריערדיגע שמועסן. דערצו: ChaosGPT, אן אויטאנאמישער אגענט וואס מען האט געגעבן דעם ציל "צו פארניכטן די מענטשהייט", איז ארויסגעווארפן געווארן פון Twitter.
- **11 אפריל:** עטליכע אנאליזן – א לערנונג אין Stochastic Reinforcement Learning מיט Approximation. LLMMaps צייגט וויזועל וואס שפראך-מאדעלן וויסן. א ארטיקל ווייזט אז א קלענערער מאדעל וואס מען טריינירט לענגער גיט אפט בעסערע רעזולטאטן, כאטש עס איז נישט די "אפטימאלע" צעטיילונג פון רעכנ-קראפט.
- **13 אפריל:** אגענטן פאר וויסנשאפט – GPT-4, מיט צוטריט צו וועב-זוכן, דאקומענטאציע, Python און אן עקספערימענט-API, האט געקענט סינטעזירן כעמישע שטאפן. אין Lawrence Berkeley לאבאראטאריע טוען ראבאטן און AI די ארבעט טאג און נאכט און זוכן נייע מאטעריאלן אליין, וואס פארשנעלערט די פארשונג ביז 100 מאל.
- **13 אפריל:** ChatGPT אין 37 שפראכן – א שטודיע אויף 7 אויפגאבן און 37 שפראכן געפינט אז אויסער ענגליש לייסט ChatGPT ערגער ווי פריערדיגע מאדעלן.
- **17 אפריל:** Jason Wei איבער גרייס – Jason Wei האט געהאלטן א פרעזענטאציע איבער ווי מאדעלן ענדערן זיך ווען זיי וואקסן, און איבער "אויפשטייענדיגע" פעאיגקייטן. לויט אים זענען נאך פאראן פעאיגקייטן וואס מען האט נאך נישט אנטדעקט.
- **17 אפריל:** אפטימיזערס – Automatic Gradient Descent באזייטיגט א טייל היפערפאראמעטערס. Meta's פארשער האלטן אז Adam איז די סיבה פון שפרינגען אינעם טריינירן. אן ארטיקל פרעגט צי Adam איז טאקע דער בעסטער, אדער ער איז נאר געווארן דער בעסטער ווייל מען טוישט אים קיינמאל נישט.
- **17 אפריל:** SVM ביים זוכן – Andrej Karpathy ווייזט ווי מען ניצט SVM אנשטאט KNN צו געפינען צופאסונגען אין עמבעדינגס.
- **17 אפריל:** דאטא-זאמלונגען – OpenAssistant: 161,443 מעסעדזשעס מיט איבער 400 טויזנט באצייכענונגען. Multimodal C4: 103M דאקומענטן, 585M בילדער און 43B טאקענס. דאטא געשאפן דורך GPT-4 פאר צופאסן LLaMA ווערט פרייגעגעבן. PMC-Patients: 167 טויזנט צוזאמענפאסונגען פון פאציענטן.
- **18 אפריל:** Cerebras-GPT – א משפחה פון מאדעלן טריינירט אויף א "רעכנ-אפטימאלן" אופן. די מאדעלן און דער קאוד זענען אפן.
- **19 אפריל:** Gist טאקענס – "Gisting" דריקט צוזאמען פראמפטס אויף ווייניגע טאקענס וואס מען קען ווידער ניצן, און שפארט אזוי אסאך רעכנ-קראפט.
- **20 אפריל:** פלאנירן און פראגראמירן מיט שפראך – Low-code LLM ניצט וויזועלע פראגראמירונג צו קאנטראלירן מאדעלן. "Learning to Program" ברענגט א 18.3% פארבעסערונג אויף AMPS. ספעציעלע פראמפטס פארבעסערן לאנגע פלענער. SCM לאזט באהאנדלען לאנגע אינפוטס מיט א זכרון-סיסטעם.

- **20 אפריל: 3D** – עטליכע ארבעטן איבער 3D: Lift3D, Nerfbusters, Farm3D (3D חיות), רעאל-צייט NeRF, VisFusion, HOSNeRF (ביז 40%-50% בעסער אויף LPIPS), TextDeformer, PoseVocab און ray conditioning.
- **21 אפריל: מעדיצינישע בילדער** – א נייע זאמלונג און אפענע מאדעלן פאר אפטיילן מעדיצינישע בילדער.
- **24 אפריל: RL פאר-טריינירן** – Sergey Levine ערקלערט ווי offline RL קען פאר-טריינירן מאדעלן וואס ארבעטן גוט שפעטער אויך online.
- **25 אפריל: ChatGPT און "אנגסט"** – פראמפטס פון קאמפיוטאציאנעלער פסיכיאטריע וואס רופן ארויס אנגסט מאכן אז ChatGPT זאל שטארק רעאגירן. ער ווערט מער פאררופן אויסצופרובירן, און אין געוויסע פעלער מער פארשטעלט.
- **25 אפריל: מאלעקולן** – SILVR שאפט נייע מאלעקולן וואס פאסן אריין אין א פראטעין, אויסגעפרובירט אויף פראגמענטן פון SARS-CoV-2.
- **26 אפריל: 3D דערקענען און self-supervised** – א ספארס סיסטעם וואס פאראייניגט בילדער מיט LiDAR איז 2.7 מאל שנעלער. דערצו א "קאכבוך" פאר self-supervised learning.
- **27 אפריל: עטיק, איבערזעצן און RL** – א בענטשמארק מעסט צי מאדעלן זוכן מאכט. S4 בלייבט 4 BLEU פונקטן הינטער Transformer ביים איבערזעצן, ביז מען לייגט צו "attention". אן ארטיקל טענהט אז RL קען העלפן קעגן "האלוצינאציעס".
- **28 אפריל: פוסבאל-ראבאט** – Deep RL האט טריינירט א קליינעם ביליגן ראבאט מיט 20 געלענקן צו שפילן 1-אויף-1 פוסבאל.
- **28 אפריל: פארשטעלטקייט** – א פארשונג-ארבעט באהאנדלט די סכנות פון פארשטעלטקייט אין מאדעלן ווי ChatGPT.

געצייג פאר פראגראמירער און אפן-סאורס

- **3 אפריל: Chroma, nanobind, Marvin** – Chroma איז אן אפן-סאורס דאטאבאזע פאר עמבעדינגס. nanobind, פונעם שאפער פון pybind11, איז א לייכטערער וועג צו פארבינדן C++ מיט Python. Marvin איז א ביבליאטעק צו בויען AI ווייכווארג.
- **4 אפריל: Auto-GPT און BabyAGI** – Auto-GPT איז אן עקספערימענטאלע אפן-סאורס אפליקאציע וואס לאזט GPT-4 ארבעטן גאנץ אויטאנאמיש, למשל צו פירן א ביזנעס. ער איז געווען #1 אויף Twitter און GitHub צוויי וואכן לאנג. מיט BabyAGI און Camel שטייגט די באגייסטערונג פאר אויטאנאמישע AI אגענטן. Latent Space טענהט אז אגענטן זענען די קומענדיגע "גרויסע אפליקאציע".
- **6 אפריל: לימודים און Vocode** – א לימוד איבער די יסודות פון שפראך-מאדעלן. Lyft ערקלערט ווי זי ניצט עמבעדינגס. Vocode בויט אפליקאציעס מיט קול. Andrej Karpathy האט א קורס צו בויען נעוראלע נעצן פון אנפאנג אן.
- **10 אפריל: llama.cpp** – Georgi Gerganov's llama.cpp לאדנט LLaMA 100 מאל שנעלער און האלבירט דעם זכרון, אזוי אז מען קען עס לויפן אויף פערזענליכע קאמפיוטערס.

- **10 אפריל:** קאוד-העלפערס – AI Shell און Engshell פארוואנדלען ענגליש אויף שעל-באפעלן. Tabby איז אן אפן-סאורס אלטערנאטיוו צו Wolverine. GitHub Copilot. GPT-4 ניצט צו פאררעכטן סקריפטס. Promptr ענדערט קאוד מיט GPT.
- **11 אפריל:** Openplayground, UpTrain, POT – Openplayground לויפט שפראך-מאדעלן אויפן לאפטאפ. UpTrain קוקט נאך ML אפליקאציעס. Python Optimal Transport איז שנעלער געווארן. א ביבליאטעק פאר טריינירן אויף TPUs, פון וועלכער עס שטאמען די FlanT5 מאדעלן.
- **12 אפריל:** TurboPilot – TurboPilot איז א Copilot וואס מען לויפט ביי זיך אליין. ער ניצט Salesforce's Codegen מיט 6 ביליאן פאראמעטערס אין 4GB RAM.
- **13 אפריל:** Consistency Models – OpenAI האט ארויסגעלאזט קאוד און טריינירטע מאדעלן פון Consistency Models, א נייער טיפ מאדעל, שנעלער ווי דיפוזיע און סטאבילער ווי GAN. דערצו א Telegram באט מיט ChatGPT און Whisper.
- **19 אפריל:** Street Fighter, Window.ai, Auto Evaluator – אן RL אגענט וואס געווינט אין א קאמף-שפיל. Window.ai איז א בלעטערער-צוגאב פאר AI מאדעלן. Auto Evaluator שאפט און ענטפערט פראגעס אויטאמאטיש.
- **21 אפריל:** Semantic Kernel – Microsoft's Semantic Kernel איז א לייכטער SDK צו פארבינדן שפראך-מאדעלן מיט געווענליכע פראגראמיר-שפראכן. עס האט פראמפט-מוסטערן, קייטן פון פונקציעס, וועקטאר-זכרון און פלאנירן.
- **24 אפריל:** W&B Prompts, ControlNet – W&B Prompts איז א געצייג פאר פראמפט-אינושנירן מיט פארבינדונגען צו LangChain און Llama Academy. OpenAI לערנט מאדעלן נייע APIs. פיר נייע ControlNet מאדעלן. א געצייג קלאנירט א קול פון א פאר סעקונדעס.
- **25 אפריל:** אינעם בלעטערער – Web LLM און WasmGPT לויפן טשעטבאטס גלייך אינעם בלעטערער. CopilotForXcode ברענגט Copilot צו Xcode.
- **26 אפריל:** HealthGPT, ResearchGPT – HealthGPT לאזט רעדן מיט דער געזונט-דאטא אין Apple Health. ResearchGPT שרייבט אנאליז-קאוד און ערקלערט די רעזולטאטן.
- **3 אפריל:** נאך אפליקאציעס – Gamma פאר פרעזענטאציעס; Rask איבערזעצט ווידיאס אין +60 שפראכן; June AI ענטפערט פראגעס איבער א פראדוקט מיט SQL; Audio2Doc; Landing AI; Apollo; Support Guy; Chefff פאר רעצעפטן; OASIS; Marketing Co-Pilot; AI Functions; Bloks; Zentask; Chirper.ai; Mage; MeetGeek, א סאציאלע נעץ פאר טשעטבאטס; SpeechGPT; Cheetah; Imagica; Chatscout; STUDIO AI.

פירמעס, געלט און ביזנעס

- **5 אפריל:** LangChain – LangChain, אן אפן-סאורס געצייג-קאסטן צו פארבינדן מאדעלן, האט אויפגענומען \$10 מיליאן.
- **5 אפריל:** Expedia, Amazon – Expedia האט אריינגעשטעלט GPT-4 פאר פלאנירן רייזעס. Amazon האט געעפנט אן "אקסעלעראטאר" פאר סטארטאפס אין גענעראטיווע AI.

- **6 אפריל:** Google TPU v4 – Google's סופער-קאמפיוטער TPU v4, מיט איבער 4,000 TPUs, איז לויט Google 1.2-1.7 מאל שנעלער ווי Nvidia's A100 און ניצט 1.3-1.9 מאל ווייניגער כח. ער ארבעט שוין זינט 2020 און מען האט אויף אים טריינירט PaLM.
- **6 אפריל:** Meta – Meta פלאנירט ביז דעצעמבער צו פארקויפן גענעראטיווע AI פאר רעקלאמעס. Mark Zuckerberg האט געזאגט אז Meta וויל ברענגען AI אגענטן צו ביליאנען באניצער: טשעט אין WhatsApp און Messenger, געצייג פאר בילדער אויף Facebook און Instagram, און שפעטער אינעם מעטאווערס.
- **7 אפריל:** וועקטאר-דאטאבאזעס – Chroma האט אויפגענומען \$18 מיליאן, צו געבן שפראך-מאדעלן "זכרון" און אזוי פארמינערן האלוצינאציעס. Fixie האט אויפגענומען \$17 מיליאן פאר אריינשטעלן מאדעלן אין פראדוקטן. Pinecone האט אויפגענומען \$100 מיליאן (Series B) פאר "לאנג-טערמיניגן זכרון" פון שפראך-מאדעלן.
- **7 אפריל:** Samsung – ארבעטער ביי Samsung האבן אין ווייניגער ווי א חודש דריי מאל ארויסגעלאזט סודות'דיגע דאטא דורך ChatGPT. Samsung Semiconductor בויט אן אייגענעם AI פאר די ארבעטער.
- **7 אפריל:** Google און Microsoft – Microsoft האט צוגעלייגט Bing טשעט צו SwiftKey אויף Android. Sundar Pichai זאגט אז Google וועט צולייגן טשעט צום זוכן. Samsung קלערט צו מאכן Bing דעם הויפט-זוכער, און Google's זוך-ביזנעס, וואס איז לעצטע יאר געווען ווערט \$162 ביליאן, ארבעט אויף "Magi". Google וועט אויך אריינשטעלן גענעראטיווע AI אין רעקלאמעס.
- **10 אפריל:** Stability AI – Stability AI ברענט דורך איר געלט און שטייט אויף שוואכע פיס.
- **11 אפריל:** Amazon – Amazon זאגט אירע ארבעטער אז זי בלייבט נישט הינטערשטעליג. Swami Sivasubramanian דערציילט וועגן איבער 100,000 ML קונים.
- **12 אפריל:** OpenAI אין יאפאן – Sam Altman קלערט צו עפענען א ביורא אין יאפאן נאך א באגעגעניש מיט Fumio Kishida. אינוועסטארן האבן אויסגערעכנט 44 סטארטאפס מיט גרויס הבטחה, אריינגערעכנט OpenAI, Stability AI, LangChain, Fixie און Cursor.
- **13 אפריל:** Elon Musk – Musk האט געקויפט בערך 10,000 GPUs פאר Twitter און דינגט AI עקספערטן. ער האט דערציילט וועגן "TruthGPT", א AI וואס זוכט "מאקסימום אמת". ער האט אפגעשניטן OpenAI's צוטריט צו Twitter דאטא, און פירט שמועסן איבער א פירמע X.AI.
- **17 אפריל:** Anthropic's קאסטן – Anthropic פלאנירט אויסצוגעבן א ביליאן דאלאר אויף א מאדעל. דער גרעסטער קאסט וועט זיין די געהאלטן פון די פארשער, נאך דעם קומט די רעכן-קראפט.
- **18 אפריל:** Scale באריכט – 56%-72% פון פירמעס האבן פארשנעלערט זייערע AI פלענער. א סך בעווארצוגן אפן-סאורס, און OpenAI's APIs פירן ביי די פארמאכטע מאדעלן.
- **19 אפריל:** Microsoft טשיפס – Microsoft ארבעט לויט באריכטן אויף אייגענע AI טשיפס, כדי נישט צו דארפן זיך פארלאזן אויף Nvidia.
- **20 אפריל:** GPT-4 אין געזונט – Microsoft און Epic ברענגען GPT-4 צו שרייבן ענטפערס פאר פאציענטן און צו אנאליזירן מעדיצינישע רעקארדס.
- **20 אפריל:** טאלאנט-קריג – טעכנאלאגיע-פירמעס ציען ארויס AI טאלאנט פון אוניווערסיטעטן.

- **21 אפריל:** Google DeepMind – Google Brain און DeepMind זענען צוזאמענגעשטאלצן געווארן אין Google DeepMind אונטער Demis Hassabis. דער ציל איז צו פארשנעלערן פארשונג, און עס איז אויסדריקליך פאר "AI פראדוקטן" - א טויש פון פריער.
- **24 אפריל:** קאסטן – לויט SemiAnalysis צאלט OpenAI ביז \$700,000 א טאג צו האלטן ChatGPT. Epoch AI שאצט אז דאס טריינירן GPT-4 האט אפשר געקאסט \$43 מיליאן.
- **27 אפריל:** Anthropic און Anthropic – Scale ארבעט צוזאמען מיט Scale פאר אויפפאסן, אויסשטעלן און זיכערהייט פון מאדעלן.
- **17 אפריל:** Princeton – Princeton האט געעפנט א צענטער פאר שפראך און אינטעליגענץ.

רעגירונגען, געזעצן און זיכערהייט

- **3 אפריל:** איטאליע און אייראפע – CAIDP האט אריינגעגעבן א קלאגע ביי דער FTC וואס קען אפשטעלן ChatGPT. איטאליע האט פארבאטן ChatGPT צוליב EU פריוואטקייט-געזעצן, און דייטשלאנד קלערט אויך. OpenAI וועט פארשלאגן לייוונגען פאר איטאליע, און די European Data Protection Board האט געשאפן א ChatGPT קאמיטעט.
- **4 אפריל:** Clearview AI – Clearview AI האט געבויט א דאטאבאזע פון 30 ביליאן בילדער פון Facebook, וואס די פאליציי האט גענוצט כמעט א מיליאן מאל.
- **5 אפריל:** אמעריקע – Biden האט זיך געטראפן מיט PCAST איבער AI. די אדמיניסטראציע קלערט רעגולאציעס, און רעגולירער ווארענען אז זיי האבן שוין די מאכט צו באקעמפן פארשטעלטקייט אין AI.
- **6 אפריל:** OpenAI זיכערהייט – OpenAI באשרייבט אירע ארבעטן פאר אמת, קינדער-זיכערהייט און פריוואטקייט. GPT-4 איז פארטיג געווען 6 חדשים פארן לאנסירן, אבער מען האט אים אפגעהאלטן צוליב זיכערהייט. OpenAI צאלט ביז \$20,000 פאר געפינען באגס.
- **6 אפריל:** קלאגעס און Snapchat – אן אויסטראלישער מעיאר סטראשעט צו קלאגן OpenAI ווייל ChatGPT האט פאלש געזאגט אז ער איז געזעסן פאר שוחד. Snapchat לייגט צו שוץ-מיטלען פאר איר טשעטבאט.
- **12 אפריל:** כינע – כינע האט ארויסגעגעבן פלאן-כללים פאר גענעראטיווע AI, וואס באשטימען למשל אויך וועלכן אינהאלט מען מעג שאפן. דאס קומט נאכדעם וואס Alibaba און Baidu האבן לאנסירט סערוויסעס.
- **17 אפריל:** אייראפעאישער פארלאמענט – דער פארלאמענט גרייט שטרענגע מאסנאמען, אריינגערעכנט אז טשעטבאט-מאכער זאלן מוזן ארויסזאגן צי זיי ניצן קאפירייט מאטעריאל.
- **4 אפריל:** קאפירייט – קלאגעס קעגן Stable Diffusion קענען שאפן א "רעש" אין געזעץ. Universal Music האט געבעטן Spotify און Apple צו פארמיידן אז AI זאל נעמען פון אירע לידער. א ליד וואס איז נאכגעמאכט אויף די קולות פון Drake און The Weeknd איז אראפגענומען געווארן נאך א טענה פון UMG, און מען ריכט זיך אויף קלאגעס.
- **25 אפריל:** בריטאניע – בריטאניע גיט £100 מיליאן (בערך \$124.5 מיליאן) פאר א Foundation Model Taskforce צו אנטוויקלען אן אייגענעם AI.

- **19 אפריל:** רעגירונגען ניצן ChatGPT – יאפאן ניצט ChatGPT צו מאכן רעגירונג-תקנות גרינגער צו פארשטיין. דער קאנגרעס האט געקויפט ChatGPT Plus 40 ליצענצן פאר זיינע ביוראס.
- **11 אפריל:** וואלן – עקספערטן ווארענען וועגן AI ביי וואלן. די RNC האט ארויסגעלאזט א רעקלאמע געשאפן מיט AI קעגן Biden.
- **27 אפריל:** Palantir און LAION – Palantir האט געוויזן א AI פאר מלחמה. דער ויכוח איבער רעגולירן LAION, מיט איבער 5 ביליאן בילדער, ווערט שארפער.

אנאליזן און מיינונגען

- **3 אפריל:** Sam Altman – צוויי פראפילן פון Sam Altman באהאנדלען זיינע געדאנקען איבער AI'ס שנעלן פארשריט. Altman זאגט אז GPT5 ווערט נאך נישט טריינירט.
- **4 אפריל:** "האלוצינאציעס" – Bloomberg טענהט אז מאדעלן "האלוצינירן" נישט - זיי טראכטן אויס. Ars Technica ערקלערט פארוואס זיי טוען דאס, און ווי מען פרובירט עס צו פאררעכטן.
- **4 אפריל:** Stanford באריכט – Stanford'ס באריכט פון 386 זייטן: AI פארשווג איז אריבער פון אוניווערסיטעטן צו פירמעס, אלטע בענטשמארקס ווערן שוואך, און דער ענערגיע-פארברויך וואקסט. דער באריכט ווארנט וועגן קאנטראל פון פירמעס.
- **5 אפריל:** Gross און Friedman – Ben Thompson שמועסט ווידער מיט Daniel Gross און Nat Friedman איבער דער AI פראדוקט-רעוואלוציע.
- **10 אפריל:** בויען אפליקאציעס – Every טענהט אז נוצליכקייט וועט קומען פון צוטריט צו ריכטיגע וויסן, למשל דורך וועקטאר-דאטאבאזעס. Chip Huyen זאגט אז עס איז גרינג צו בויען עפעס שיינס, אבער שווער עס צו ברענגען צו פראדוקציע.
- **10 אפריל:** גיכקייט קעגן פארזיכטיגקייט – Microsoft און Google קלייבן גיכקייט איבער פארזיכטיגקייט. ארבעטער ביי Google האבן אנגערופן Bard "א פאטאלאגישער ליגנער", און מען האט אים פונדעסטוועגן לאנסירט.
- **13 אפריל:** דער "Llama עפעקט" – ווי Llama'ס דורכגעזיקערט האט אויסגעלעזט א פלייץ פון אפענע ChatGPT-ענלעכע מאדעלן.
- **13 אפריל:** ארבעט – א ארטיקל באטראכט די עקאנאמישע השפעה פון GPT-4. Kent Beck זאגט אז 90% פון זיינע פעאיקייטן זענען יעצט ווערט \$0, אבער די איבריגע 10% זענען ווערט מער.
- **14 אפריל:** קאסטן פאר מענטש און סביבה – א ארטיקל ווארנט וועגן ענערגיע, קוילן-שטאף אויסשטויסונג, און שאדן פאר שעפערישע פאכן.
- **20 אפריל:** John Schulman – OpenAI זאגט אז אמת איז איינע פון די גרעסטע פראבלעמען, און RLHF איז א וויכטיגער וועג עס צו פאררעכטן.
- **21 אפריל:** טשיפס און כינע – א ארטיקל באהאנדלט דעם מארק פון טשיפס, אויפן הינטערגרונט פון די שפאנונגען צווישן אמעריקע און כינע.

- **25 אפריל:** AI פאר ארבעטער – א NBER שטודיע: א AI העלפער פאר קונה-סערוויס ברענגט בערך 14% מער פראדוקטיווקייט, מערסטנס פאר ווייניגער ערפארענע ארבעטער, און פארבעסערט צופרידנקייט און דאס בלייבן אין דער ארבעט.
- **25 אפריל:** "Doomers" – Techdirt טענהט אז אנשטאט אן אויסגעלאסענעם AI האט מען אן אויסגעלאסענע פאניק.
- **19 אפריל:** פראמפטס – עצות ווי פראדוקט-מענדזשערס קענען ניצן ChatGPT; Xu Hao ווייזט ווי מען באקומט קאוד וואס טעסט זיך אליין; Mitchell Hashimoto פארגלייכט פראמפט-אינזשינירן מיט "בלינדע פראמפטס".
- **26 אפריל:** Spotify און Pew – דער CEO Spotify רופט AI "קול און שרעקליך". Pew געפינט געמישטע מיינונגען ביי אמעריקאנער איבער AI אין דינגען ארבעטער.

ווייטערע ידיעות

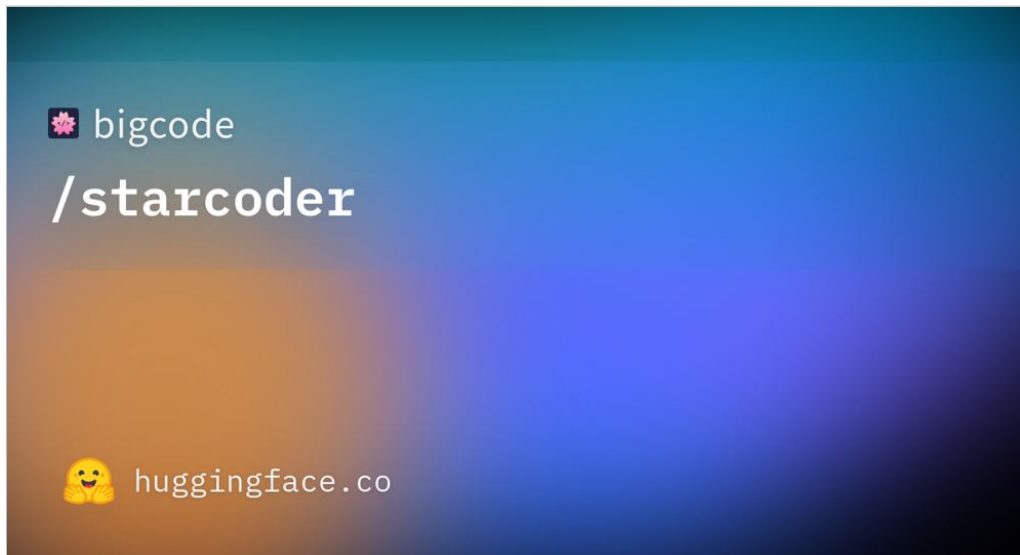
- **3 אפריל:** ChatGPT ווי מדריך – מען קלערט צי ChatGPT קען ערזעצן א מדריך פאר יונגע ארבעטער. עקספערטן ווארענען אבער זיך נישט צו פארלאזן צופיל דערויף.
- **3 אפריל:** מוח – AI העלפט פארשטיין ווי נעוראנען רעדן איינס מיטן אנדערן. Cornell האט געבויט ברילן וואס לייענען שטילע רייד. א יאפאנעזער שטודיע איבער "לייענען מחשבות" הייבט אויף עטישע פראגעס.
- **4 אפריל:** שפילן – FIFA האט לאנסירט AI World Cup AI League. קען איבערקערן די ווידיאו-שפיל אינדוסטריע.
- **4 אפריל:** ענזיימען – א AI געצייג זאגט פאראויס די פונקציע פון ענזיימען בעסער ווי די פירנדע געצייג.
- **4 אפריל:** פראמפט-אינזשינירן – גוטע פראמפט-אינזשינירן פארדינען ביז \$335,000 א יאר. 91% פון ארבעטגעבער זוכן ארבעטער מיט ChatGPT ערפארונג.
- **6 אפריל:** AI ווידיאוס – AI ווידיאוס ווערן פאפולער. עקספערטן זאגן אז זיי וועלן נישט ערזעצן מענטשן אין פילמען נאך אזוי גיך.
- **7 אפריל:** AI ארבעט און Nvidia GTC – א ארטיקל באשרייבט ווי עס פילט זיך צו ארבעטן אין AI, צווישן פארשריט און זיכערהייט. א צוזאמענפאסונג פון Nvidia'ס GTC אנאנסירונגען: NeMo, Picasso, BioNeMo און Omniverse.
- **12 אפריל:** Reddit – Reddit מאדעראטארן גרייטן זיך צו א פלייץ פון ספאם געשאפן מיט ChatGPT.
- **14 אפריל:** געצייכנטע בילדער – Meta האט געעפנט איר פראיעקט וואס באלעבט קינדער-צייכענונגען, מיט איבער 100 טויזנט ביישפילן.
- **17 אפריל:** מאדעלן ווערן שוואכער – 91% פון ML מאדעלן ווערן מיט דער צייט ערגער.
- **18 אפריל:** DuckDuckGo – DuckDuckGo ניצט AI צו צוזאמענפאסן אינפארמאציע פון Wikipedia.
- **19 אפריל:** פאטא-פרייז – Boris Eldagsen האט אפגעזאגט א פרייז פון Sony World Photography, ווייל זיין בילד איז געשאפן געווארן מיט AI.
- **21 אפריל:** לאגס – ChatGPT איז גוט אין געפינען און פאררעכטן פראבלעמען אין לאגס.

- **24 אפריל: Humane – Humane**, א פירמע פון געוועזענע Apple דיזינערס, ווייזט א מכשיר וואס פראיעקטירט טעלעפאן-רופן אויף דער האנט.
- **26 אפריל: ראבאטן אין געשעפטן** – אויטאנאמישע ראבאטן פאר זאפאס קענען שפארן ביליאנען פאר געשעפטן.
- **27 אפריל: פריילענסערס און מוזיק – AI** טרייסלט די פריילענס וועלט. טויזנטער מוזיק-פראדוצענטן שאפן AI מוזיק אויף TikTok. Discord. פרובירט AI אוואטארן.
- **28 אפריל: OpenAI אינטערוויו** – OpenAI'ס CTO רעדט איבער GPT-4, ChatGPT, און DALL-E, און איבער פאראנטווארטליכן באנוץ. א ארטיקל ווייזט וועלכע וועבזייטלעך מען ניצט צו טריינירן מאדעלן. פארשונג-ארבעטן ביי CHI 2023.

מאי 2023 איז געווען א חודש פון גרויסע לאנסירונגען און שטארקע דעבאטעס איבער סכנות און רעגולירונג. Google האט ארויסגעלאזט PaLM 2 און דערמיט פארבעסערט Bard, Colab, און אנגעהויבן אריינשטעלן AI אין זוכן. Anthropic האט אויסגעברייטערט Claude ביז 100 טויזנט טאקענס און אויפגענומען 450 מיליאן דאלאר. אויף דער אפן-סאורס פראנט זענען ארויס StarCoder, QLoRA, RWKV און א רייע אנדערע מאדעלן, באגלייט מיט א געליקטן Google דאקומענט וואס זאגט אז די אפענע קהילה כאפט אן די גרויסע פירמעס. Sam Altman האט געזאגט עדות פארן קאנגרעס. Geoffrey Hinton האט פארלאזט Google כדי צו רעדן אפן וועגן סכנות. דער חודש האט זיך פארענדיקט מיט א דערקלערונג פון פארשער אז די סכנה פון אויסשטארבן צוליב AI דארף באהאנדלט ווערן ווי מגיפות און נוקלעארע מלחמה, און מיט Nvidia's אקציע וואס איז געשפרונגען 28%.

מאדעלן און נייע לאנסירונגען

- **1 מיי: Microsoft Designer** – Microsoft האט ארויס א פריווי פון איר Designer געצייג. עס שלאגט אויטאמאטיש פאר דיזיין-עלעמענטן און אויסשטעלונג-אפציעס.
- **1 מיי: Stability AI – StableVicuna** – Stability AI האט ארויסגעלאזט א טשעטבאט וואס פירט זיך ווי ChatGPT און איז מערסטנס אפן-סאורס. ער איז אבער נישט אזוי שטארק, און עס זענען דא באגרענעצונגען אויף געשעפטליכן באנוץ.
- **2 מיי: Stability AI – DeepFloyd IF** – Stability AI האט ארויסגעלאזט DeepFloyd IF, א מאדעל וואס פארוואנדלט טעקסט אין דעטאלירטע בילדער.
- **4 מיי: Pi פון Inflection AI** – Inflection AI איז א סטארטאפ פון געוועזענע DeepMind פירער. אונטער דעם CEO Mustafa Suleyman האט זי לאנסירט Pi, א טשעטבאט וואס געדענקט פריערדיגע שמועסן און איז געבויט צו זיין אן "אקטיווער צוהערער". שפעטער זאל ער אויך העלפן ארגאנעזירן א לוח-זמנים און זיך צוגרייטן צו מיטינגען. די פירמע האט אנגעשטעלט איבער 600 טייל-צייט "לערער", אריינגערעכנט טעראפיסטן, כדי צו טריינירן דעם מאדעל ווי א קאוטש און א געטרייער באגלייטער.
- **5 מיי: StarCoder** – Hugging Face און ServiceNow האבן ארויסגעלאזט StarCoder, א פרייער מאדעל וואס שרייבט קאוד אין 86 פראגראמיר-שפראכן. עס איז דער ערשטער אפענער קאוד-מאדעל וואס דערגרייכט איבער 1@40% pass אויפן שווערן HumanEval בענטשמארק. ער קען אויסשטעלן קאוד אין מיטן און פירן א דיאלאג. א באזונדערע ארטיקל ווייזט ווי אזוי אים צו מאכן אין א טשעט-געהילף. מען קען אים שוין לויפן אויף א געווענליכן קאמפיוטער דורך GGML מיט 4-בייט קוואנטיזאציע.



2023 מ"י 5 · StarCoder



2023 מ"י 5 · StarCoder

- **5 מ"י: Replit קאוד-מאדעל** – Replit האט ארויסגעלאזט אן אפענעם קאוד-מאדעל, געטריינירט דורך געוועזענע Google פארשער אויף דער MosaicML פלאטפארמע. פאר זיין גרייס איז ער איבערראשנד שטארק אויף HumanEval. אין אן אינטערוויו האט Replit'ס AI-הויפט Reza Shabani דערציילט ווי די פירמע האט געטריינירט אן אייגענעם מאדעל פאר אירע 22 מיליאן פראגראמירער [23050406].
- **8 מ"י: MosaicML 65k קאנטעקסט** – MosaicML האט ארויסגעלאזט א מאדעל וואס קען ארבעטן מיט 65 טויזנט טאקענס קאנטעקסט, אין פארגלייך צו 8 טויזנט ביי ChatGPT. דאס איז גענוג אריינצולייגן דאס גאנצע בוך "The Great Gatsby" און לאזן דעם מאדעל שרייבן א המשך.
- **8 מ"י: Shap-E** – OpenAI ארבעט שטילערהייט אויף טעקסט-צו-3D. דער נייער Shap-E מאדעל ארבעט גוט ביי געווענליכע חפצים. ער קען אויך מאכן 3D פון 2D בילדער, נאר דארט ארבעט עס ווייניגער גוט.
- **8 מ"י: GPT-4-32k** – עס זעט אויס אז OpenAI הייבט אן צו פארשפרייטן GPT-4 מיט 32 טויזנט טאקענס קאנטעקסט.
- **9 מ"י: PaLM 2 און Bard** – Google האט אויף איר יערליכער קאנפערענץ פארגעשטעלט PaLM 2. דער מאדעל איז געטריינירט אויף כמעט פינף מאל מער דאטא ווי זיין פארגייער פון 2022, און ער איז בעסער

אין קאוד, מאטעמאטיק, שרייבן און פילע שפראכן. לויטן טעכנישן באריכט איז ער שנעלער און עפעקטיווער פון PaLM. Bard איז אריבער אויף PaLM 2 און איז יעצט צוגענגליך אין איבער 180 לענדער, אויף יאפאניש און קארעאניש, מיט 40 שפראכן וואס קומען באלד. ער באקומט אויך פארבינדונג מיט Gmail און Google Docs, וויזועלע ענטפערס און הילף ביים זוכן פעלערן אין קאוד.



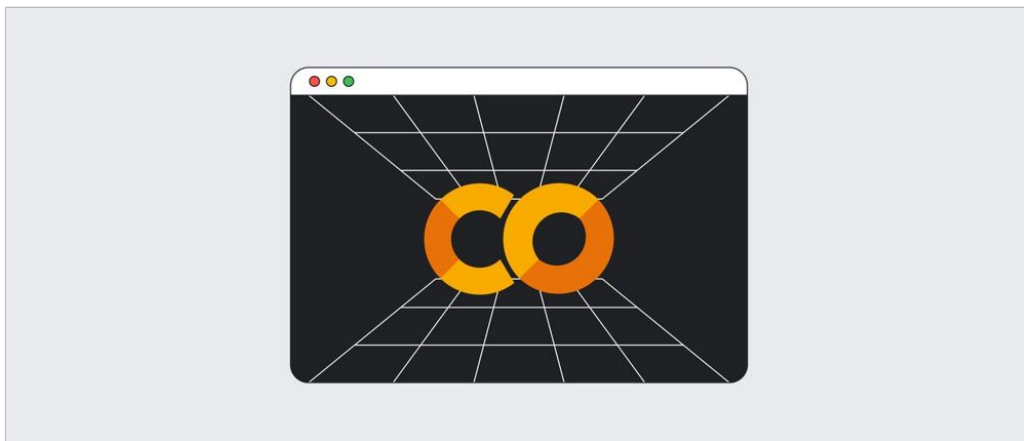
2023 און PaLM 2 · 9 מיי 2023

- **10 מיי: ImageBind** – Meta האט ארויסגעלאזט אן אפן-סאורס מאדעל וואס פארבינדט זעקס סארטן דאטא: טעקסט, קול, בילד, טעמפעראטור און באוועגונג. אלע ווערן געשטעלט אין איין געמיינזאמען "עמבעדינג" רויס. אן אנאליז לויבט די ווערסאטילקייט, אבער באמערקט אז עס קומען אמאל ארויס אומפאסיגע רעזולטאטן.
- **10 מיי: Claude's קאנסטיטוציע** – Anthropic האט דערקלערט אז Claude ווערט געטריינירט מיט "Constitutional AI". דאס הייסט אז דער מאדעל באקומט קלארע ווערטן פון א רשימה פרינציפן, אריינגערעכנט די UN דעקלאראציע פון מענטשנרעכט, אנשטאט פארבארגענע ווערטן פון RLHF.
- **10 מיי: Microsoft 365 Copilot** – Microsoft עפנט ברייטער דעם פריווי פון Copilot 365, וואס איז געבויט אויף GPT-4. דערצו קומט א נייער אינדעקס-געצייג פאר אינערליכע פירמע-דאטא און נייע פונקציעס אין Outlook, PowerPoint און Whiteboard.
- **11 מיי: MusicLM** – Google האט ארויסגעלאזט MusicLM אין דער AI Test Kitchen אפליקאציע. עס מאכט עטליכע ווערסיעס פון מוזיק פון א טעקסט-באשרייבונג.
- **11 מיי: Hugging Face Agents** – Hugging Face האט א נייעם אינטערפעיס פאר AI אגענטן, דאס הייסט מאדעלן וואס ניצן געצייג, מיט א צוזאמגעשטעלטע רשימה פון מאדעלן. StarCoder מאכט זיי באדייטנד מער ניצלעך.
- **12 מיי: Claude 100k** – Anthropic's Claude קען יעצט ארבעטן מיט 100 טויזנט טאקענס. אין א ביישפיל האט מען אריינגעגעבן 240 זייטן קאוד-דאקומענטאציע און געשטעלט פראגעס דערויף. די טעכנישע פרטים זענען נישט געטיילט געווארן.

• **12 מיי: Stability AI – Stability** האט ארויסגעלאזט אן SDK צו מאכן אנימאציעס מיט Stable Diffusion. מען קען אנהייבן פון גארנישט, פון א בילד אדער פון א ווידעא. עס קאסט אבער שווער מיט רעכנן-קראפט.

• **15 מיי: ChatGPT פלאגינס** – OpenAI לאזט ארויס איבער 70 פלאגינס פאר באצאלטע באניצער, אריינגערעכנט צוטריט צום אינטערנעט. שפעטער האבן די פלאגינס אויך באקומען שטיצע פאר Postgres און Supabase דאטאבאזעס [23052611].

• **15 מיי: Codey און Colab AI** – Google האט פארגעשטעלט Codey, א פאמיליע קאוד-מאדעלן אויף PaLM 2 וואס שלאגט פאר קאוד און טרעפט פעלערן. Colab וועט באלד באקומען קאוד-פארענדיקונג, קאוד פון געווענליכער שפראך און א טשעטבאט, אלעס צוגעפאסט פאר Python.



Codey און Colab AI · 15 מיי 2023

• **15 מיי: InstructBLIP** – מען האט טריינירט Salesforce's BLIP צו פאלגן אנווייזונגען. דאס האט שטארק פארבעסערט די רעזולטאטן, און אין געוויסע פעלער איז עס בעסער פון די פארעפנטליכטע נומערן פון GPT-4.

• **18 מיי: StableStudio** – Stability AI האט ארויסגעלאזט StableStudio, אן אפן-סאורס ווערסיע פון איר DreamStudio אפליקאציע.

• **18 מיי: Zapier AI** – Zapier האט לאנסירט א ChatGPT פלאגין און א Natural Language Actions API. זי האט אויך געמאלדן צוויי נייע AI בעטא-פונקציעס פאר אויטאמאציע.

• **19 מיי: ChatGPT אויף iOS** – OpenAI האט ארויסגעלאזט אן אפליקאציע פאר Apple. אויפן ערשטן טאג איז זי געווען די צווייטע מערסט-אראפגעלאדענע פרייע אפליקאציע.

• **22 מיי: Meta טשיפ** – Meta האט אן אייגענעם 7 נאנאמעטער טשיפ פאר AI אינפערענץ. ער פארברויכט ווייניגער קראפט און איז גלייך שנעל ווי אנדערע הארדוועיר.

• **23 מיי: Aurora GenAI** – Intel האט געמאלדן Aurora GenAI, א מאדעל מיט 1 טריליאן פאראמעטערס פאר שמועס, איבערזעצונג, צוזאמפאסונג און קאוד, אין מיטארבעט מיט אקאדעמישע אינסטיטוציעס.

• **23 מיי: MMS** – Meta's פראיעקט Massively Multilingual Speech קען דערקענען און רעדן איבער 1,100 שפראכן. דער ציל איז צו מאכן אינפארמאציע מער צוגענגליך און צו העלפן אפהיטן שפראכן וואס זענען אין סכנה צו פארשווינדן.

- **24 מיי:** Bing אין Windows – Bing Chat און Copilot-סטיל פלאגיןס, סיי פון Microsoft און סיי פון דרויסנדיגע פירמעס, קומען אריין גלייך אין Windows.
- **25 מיי:** Megabyte – Meta AI האט פארגעשלאגן Megabyte, אן ארכיטעקטור וואס ארבעט גלייך מיט בייטס אן א טאקעניזער. זי קען שאפן איבער א מיליאן טאקענס אין פארשידענע פארמאטן. דאס רעכענען גייט פאראלעל, און לויט Meta איז עס עפעקטיווער ווי טראנספארמערס.
- **25 מיי:** Photoshop AI – Adobe לייגט צו גענעראטיווע AI אין Photoshop, כדי אז אויך נישט-געלערנטע באניצער זאלן קענען ארבעטן דערמיט.
- **26 מיי:** AI אין Google זוכן – Google הייבט אן ארויסצולאזן "Search Generative Experience", AI-ענטפערס אין די זוך-רעזולטאטן. דאס איז א טייל פון א ברייטערן פלאן צו טוישן ווי די רעזולטאטן ווערן געוויזן [23050816,23051109].
- **31 מיי:** שערן שמועסן – ChatGPT האט א נייע פונקציע וואס גיט יעדן שמועס אן אייגענעם לינק צום שיקן אנדערע, אנשטאט צו שיקן סקרינשאטס.
- **31 מיי:** Naver HyperClova X – די קארעאנישע פירמע Naver פלאנט צו פארקויפן פרעמדע רעגירונגען צוגעפאסטע ווערסיעס פון איר מאדעל, לויט זייערע פאליטישע און קולטורעלע באדערפענישן. זי וועט אויך לאנסירן HyperClova X פאר די היימישע סערוויסעס.

פארשונג און טעכניק

- **1 מיי:** פארשונג אנפאנג חודש – א מעטאד שטעלט אריין מענטשן אין בילדער אויף א נאטירליכן אופן מיט א דיפיוזשן מאדעל. InstructCTG קאנטראלירט טעקסט-שאפונג דורך אנווייזונגען אין געווענליכער שפראך, אן נאכאמאל טריינירן. ChatVideo איז א סיסטעם וואס פארשטייט ווידעאס. LLaMA-Adapter V2 איז א ביליגע מעטאד וואס לערנט LLaMA צו פאלגן אנווייזונגען און צו פארשטיין בילדער.
- **2 מיי:** דיפיוזשן און סינטעטישע דאטא – א פארשונג-ארבעט שטעלט באגרענעצונגען אויף וואס א דיפיוזשן מאדעל טאר ארויסגעבן. א צווייטע ווייזט אז סינטעטישע בילדער פון דיפיוזשן מאדעלן פארבעסערן באדייטנד די קלאסיפיקאציע אויף ImageNet. א דריטע געפינט אז אויסקלייבן ענליכע בילדער פון די טריינינג-דאטא גיט נאך בעסערע רעזולטאטן ווי צופאסן דעם מאדעל.
- **2 מיי:** LLaMA פאר מעדיצין – LLaMA ווערט מער ווי 30% בעסער אויף מעדיצינישע אויפגאבן ווען מען טריינירט אים ווייטער אויף מעדיצינישע דאטא.
- **2 מיי:** קול – Microsoft's NaturalSpeech 2 פראיעקט ארבעט צו פארבעסערן קול-טעכנאלאגיע. א ווערסיע פון Whisper אויף Jax איז 70 מאל שנעלער און שרייבט אראפ א שעה רעדן אין עטליכע סעקונדעס.
- **2 מיי:** בילדער פונעם מוח – וויסנשאפטלער ניצן AI צו דעקאדירן בילדער גלייך פון מוח-סקענס, וואס קען העלפן פארשטיין ווי דער מוח פארארבעט וואס ער זעט.
- **3 מיי:** וועלכע ביכער געדענקט GPT – א פארשונג-ארבעט געפינט אז די מאדעלן האבן כמעט אינגאנצן אויסווייניג געלערנט אסאך פאפולערע ביכער און אפילו עטליכע ווייניג-באקאנטע.

- **3 מיי: אוואטארן און 3D** – עס זענען ארויס מעטאדן צו שאפן די הויט-טעקסטור פון א 3D אוואטאר פון איין בילד, StyleAvatar פאר רעאליסטישע פארטרעט-אוואטארן אין רעאלע צייט פון איין ווידעא, און 3D נאטור-סצענעס פון איין ביישפיל.
- **3 מיי: Stable Diffusion פאר \$47k** – MosaicML האט איבערגעטריינירט Stable Diffusion פאר 47 טויזנט דאלאר אויף פובליק קלאוד און ארויסגעגעבן דעם קאוד. באניצער האבן געשאצט דעם רעזולטאט העכער ווי Stable Diffusion 2.0.
- **4 מיי: Unlimiformer** – א מעטאד וואס לאזט יעדן טראנספארמער מאדעל ארבעטן מיט טעקסט אן א לענג-גרעניץ. עס קען צוזאמפאסן גאנצע ביכער.
- **4 מיי: GPT-4 און סיבה** – OpenAI'ס מאדעלן זענען איבערראשנד גוט אין טרעפן סיבה און תוצאה. זיי מאכן אבער אויך מאדנע טעותים.
- **4 מיי: קליינע מאדעלן** – אן אנאליז פרעגט צי קליינע מאדעלן קענען אנכאפן די גרויסע אין שכל, ווען מען ניצט טריינירן אויף קאוד, chain-of-thought און RL. א פארשווג-ארבעט געפינט אז "דיסטילירן" א גרויסן מאדעל אין א קלענערן איז אפט ביליגער ווי אנצייכענען נאך דאטא.
- **4 מיי: Pick-a-Pic** – אן אפענע סאמלונג פון וועלכע בילדער באניצער האבן בעסער ליב, און PickScore, א מאדעל וואס זאגט פאראויס וואס מענטשן וועלן אויסקלייבן.
- **5 מיי: RLHF דערקלערט** – Chip Huyen'ס ארטיקל דערקלערט אויספירליך ווי אזוי מען לערנט א מאדעל לויט מענטשליכע באמערקונגען (RLHF).
- **5 מיי: קאנטראל און טראנספער** – עס זענען ארויס מעטאדן צו קאנטראלירן בילד-שאפונג דורך פאנט-פארב און גרייס אין דעם טעקסט, און VPGTrans, וואס מאכט ביליגער צו בויען ראיה-שפראך מאדעלן.
- **8 מיי: AI אין שווערע שמועסן** – ווען א טשעטבאט איז דער פארמיטלער אין שמועסן איבער צעטיילנדיגע טעמעס, באריכטן ביידע צדדים בעסערע רעזולטאטן און מער פארשטענדעניש.
- **8 מיי: ווייטערע ארבעטן** – ZipIt פאראייניגט מאדעלן וואס זענען געטריינירט אויף פארשידענע אויפגאבן, ביז 60% בעסער ווי פריערדיגע מעטאדן. HiPool ניצט גראפן פאר לאנגע דאקומענטן. VERA שאצט אפ צי א זאץ איז לויטן שכל הישר מעגליך. Otter און MIMIC-IT זענען פאר מולטימאדאלע אנווייזונגען.
- **8 מיי: scGPT** – אן אפענער טראנספארמער פאר ביאלאגיע פון איינצלנע צעלן. מען קען אראפלאדן די וואגן.
- **9 מיי: בענטשמארקס און מעטאדן** – Cola איז א בענטשמארק פאר צוזאמשטעלונג פון חפצים מיט אייגנשאפטן. WikiWeb2M האט 2 מיליאן וועב-זייטלעך. א נייע מעטאד בויט decision trees מיט gradient descent.
- **10 מיי: פארשוונג 10טן מאי** – OpenAI האט געוויזן אז א שפראך-מאדעל קען דערקלערן וואס יחידישע "נייראנען" אין אן אנדערן מאדעל טוען. ווען דאס ווארט "דאלאר" ווערט פאראויסגעזאגט, ווערן נייראנען פון געלט און אמעריקע אקטיוו. FrugalGPT ניצט א קייט פון מאדעלן צו צאלן ווייניגער פאר API, מיט גלייכע אדער בעסערע רעזולטאטן. FedIT לערנט מאדעלן פון פילע מאשינען אן צו שאטן פריוואטקייט. iChat צוזאמען מיט SUR-adapter פארבעסערן ארבעט מיט בילדער.

- **11 מיי:** **ביליגער און שנעלער** – לויטן deeplearning.ai בריוו קען GPT-3.5-turbo שאפן גענוג טעקסט פאר א שעה לייענען פאר בלויז 8 סענט. אריינשטעלן אלגאריטמען אין מאדעלן פארבעסערט ענטפערס מיט ראיות מיט BERT. 6.4% קען מען טריינירן אן attention, מיט state space מאדעלן.
- **12 מיי:** **3D און ווידעא** – עס זענען ארויס אן איבערבליק פון טעקסט-צו-3D, ווידעא פון סקיצעס מיט ControlNet, א 3D פנים פון איין בילד, און VideoChat.
- **15 מיי:** **בילדער** – א מעטאד מאכט פארשוואומענע בילדער קלאר מיט עקזיסטירנדע טעקסט-צו-בילד מאדעלן, און Vicuna טרעפט חפצים אין א בילד דורך פראגעס.
- **16 מיי:** **סטירן און ראיה** – מען קען "סטירן" א מאדעל, אים מאכן למשל פרייליכער, דורך צולייגן אקטיוואציע-וועקטארן. עס זענען אויך ארויס ווידעא פראגע-ענטפער מיט א קייט מאדעלן, EfficientViT, און Nvidia's o'NFL וואס בויט 3D סצענעס פון LiDAR.
- **17 מיי:** **TinyStories און SuperICL** – גאר קליינע מאדעלן, אפילו מיט איין בלאק, קענען שרייבן ריכטיגע מעשהלעך ווען מען באגרענעצט דעם ווערטער-אוצר. SuperICL פארבינדט גרויסע מאדעלן מיט קליינע. MMG-Ego4D איז א בענטשמארק פאר דערקענען אקציעס.
- **18 מיי:** **DarkBERT און סימבאלן** – DarkBERT איז געטריינירט אויפן Dark Web. ביי Symbol Tuning פארטוישט מען לעיבלס מיט צופעליגע סימבאלן, וואס פארבעסערט שכל ביי נייע אויפגאבן. Boundless DAS ווייזט ווי Alpaca לייזט א נומער-פראבלעם.
- **19 מיי:** **FastComposer און אנדערע** – FastComposer מאכט פערזענליכע בילדער מיט פיל פערזאנען אן נאכטריינירן. אנדערע ארבעטן פארשן צי מאדעלן קענען זיך לערנען צו פארהאנדלען, און ווייזן אז שפראכן אויסער ענגליש ווערן טייערער צעטיילט אין טאקענס.
- **22 מיי:** **DragGAN און מער** – DragGAN לאזט טוישן א בילד דורך שלעפן פונקטן אויף אן ארט. ווייטער זענען ארויס mLongT5 פאר לאנגע טעקסטן אין פילע שפראכן, טריינירן אין א סימולאטאר וואס מאכט קליינע מאדעלן גלייך צו ChatGPT, און ONE-PEACE פאר בילד, קול און שפראך.
- **22 מיי:** **QLoRA – QLoRA**, מיט 4-ביט NormalFloat אין Hugging Face, לאזט טריינירן א מאדעל מיט 65 ביליאן פאראמעטערס אויף א GPU מיט 48GB.
- **23 מיי:** **דיפיוזשן מיט RL** – מען טריינירט דיפיוזשן מאדעלן מיט RL, למשל מיט LLaVA ווי ריכטער אדער מיט DDPO, אן נאך דאטא אדער מענטשליכע אנצייכענונגען.
- **23 מיי:** **RLHF, CRITIC, CoDi** – אן ארטיקל טענהט אז RLHF איז נישט קיין כישוף-שטעקן פאר "אליינמענט". CRITIC לאזט א מאדעל פרובירן און פארריכטן אליין זיינע ענטפערס מיט געצייג. CoDi שאפט יעדע קאמבינאציע פון טעקסט, בילד, ווידעא און קול.
- **24 מיי:** **RWKV** – די לאנג-ערווארטעטע פארשונג-ארבעט איבער RWKV איז ארויס. עס פארבינדט טריינירן אויף פאראלעל ווי טראנספארמערס מיט די עפעקטיווקייט פון RNN. עס איז שוין אריין אין Hugging Face Transformers.
- **24 מיי:** **Safetensors און VideoLLM** – EleutherAI באשטעטיגט אז Safetensors שיצט פון לויפן שעדליכן קאוד ווען מען לאדט א מאדעל. VideoLLM ניצט שפראך-מאדעלן צו פארשטיין ווידעאס.

- **25 מיי: SEAHORSE** – עס זענען ארויס 96 טויזנט צוזאמפאסונגען אין פילע שפראכן, אפגעשאצט דורך מענטשן אויף זעקס זייטן, און meta-in-context learning.
- **26 מיי: Goat און TWIST** – Goat איז א קליינער LLaMA מאדעל וואס שלאגט GPT-4 אין חשבון, דאנק דעם טאקענייזער וואס צעטיילט נומערן אויף איינצלע ציפערן. RecurrentGPT שרייבט טעקסט פון יעדע לענג. TWIST איז דער גרעסטער רייד-מאדעל ביז יעצט.
- **29 מיי: Voyager און דאטאסעטס** – Voyager לאזט GPT-4 זיך פארבעסערן אין Minecraft דורך שרייבן קאוד, אן טריינירן. ווייטער: פארריכטן טעקסט-צו-SQL, באמערקונגען פון באניצער, און TheoremQA מיט 800 פראגעס אויף 350 טעארעמען.
- **30 מיי: פאקטן** – דריי ארבעטן: איין מאדעל "פארהערט" דעם אנדערן ווי אין געריכט. "according-to" פראמפטן מיט דעם QUIP-Score. אויך ווי מאדעלן רעאגירן אויף סותר'דיגע אינפארמאציע.
- **31 מיי: סוף חודש** – MIT פארשער לאזן עטליכע טשעטבאטס דעבאטירן כדי צו פארמינערן "האלוצינאציעס", אבער פארשער זארגן אז דער פראבלעם איז אפשר נישט צו פארריכטן. Gorilla איז בעסער פון GPT-4 אין שרייבן API רופן (APIBench). אנדערע: GPT-4 איז נישט קיין יושר'דיגער ריכטער אן פארריכטונגען, GILL, ProlificDreamer און ControlVideo.

געצייג פאר פראגראמירער און אפן-סאורס

- **1 מיי: געצייג** – Wonnx (אין Rust פארן וועב), Guidde, און GPT4Free.
- **2 מיי: געצייג** – Godot-Dodo, H2OGPT, NewsBytes, Landing AI.
- **3 מיי: Mojo און מער** – Mojo, א נייע שפראך וואס פארבינדט Python מיט C. Lamini, און Mr. Ranedeer, א לערער-פערזענלעכקייט פאר GPT-4.
- **4 מיי: OpenLLaMA און קורסן** – OpenLLaMA, אן אפענע איבערבוינג פון LLaMA. צוויי קורסן פון Databricks. Chat Chat און folk.
- **5 מיי: געצייג** – Databerry, DeepDoctection, נאטעבוק-מאניטארינג, AudioPen.
- **8 מיי: Kadoa** – אויסציען דאטא פון וועבסייטס מיט AI.
- **9 מיי: MLC LLM און PandasAI** – MLC LLM לויפט מאדעלן אויף יעדע הארדוועיר. ווייטער: PandasAI, RasaGPT, Metering.ai.
- **10 מיי: געצייג** – Feathery, AutoPR.
- **11 מיי: PrivateGPT** – PrivateGPT, Chain-of-Thought-Hub, LLM Chain אין Rust, Raycast Pro.
- **12 מיי: HuggingChat UI** – Hugging Face האט געעפנט דעם קאוד פון HuggingChat. ווייטער: Metabob, Flowise.
- **15 מיי: געצייג** – Open-LLaMa, PaLM (אפענע אימפלעמענטאציע), Jax, Cohere LLM University, shard_map, LangChain retrieval, Bard API.
- **16 מיי: Microsoft Guidance** – Guidance, א פראמפט-שפראך פון Microsoft. GPT4Tools, Databerry.ai.

- **17 מיי:** געצייג - Dify, LLM אין בראווער דורך Metaphor, WebGPU, LangChain, ראבאטיק-קוואלן, Nexus.
- **18 מיי:** געצייג - "Numbers Every LLM Dev Should Know", Smol-Developer.
- **19 מיי:** געצייג - Smallville, Microsoft פארגלייך ChatGPT מיט אפן-סאורס, Betafi.
- **22 מיי:** געצייג - SuperAgent, Puddl, Foqus.
- **23 מיי:** געצייג - GPT-JSON, Zeda.io.
- **24 מיי:** געצייג - DB-GPT, Scikit-LLM.
- **25 מיי:** געצייג - ChainForge, LoopGPT, ControlVideo, Desku.
- **26 מיי:** געצייג - a16z AI Canon, Gorilla, Lit-Parrot, Tyles.
- **29 מיי:** געצייג - Anse, ClipnoteAI, LangChain, HireAI, GDrive.
- **30 מיי:** **Sophia און Godot** - Sophia אפטימיזער. מאדעלן געטריינירט פאר Godot שלאגן. GPT-4. Docta. Samantha.
- **31 מיי:** געצייג - Threestudio, Freepik, CapeChat.

פירמעס, געלט און ביזנעס

- **1 מיי:** **Amazon** - Amazon בויט א מאדעל פאר Alexa און וועט ברענגען ChatGPT-סטיל AI אין Echo. Amazon'ס קלאוד קונים ווארטן נאך אויף צוטריט, וואס ברענגט זארג אז דער געצייג איז נאך נישט פארטיג.
- **2 מיי:** **Microsoft** - Edge האט Bing AI. פלאנט א פריוואטע ChatGPT פאר פירמעס.
- **3 מיי:** **DGX H100** - Nvidia'ס DGX מיט H100 טשיפס הייבט אן צו ווערן געשיקט.
- **3 מיי:** **פארבאטן** - Samsung און Apple באגרענעצן אנגעשטעלטע פון ChatGPT, אויס מורא פאר דאטא-ליקן.
- **4 מיי:** **Google Checks** - א פלאטפארמע פאר פריוואטקייט-פארשריפטן.
- **4 מיי:** **Chegg** - די אקציע איז געפאלן 40% נאכדעם וואס די פירמע האט געזאגט אז ChatGPT שאדט איר.
- **5 מיי:** **Google דאקומענט** - א געליקטער דאקומענט טענהט אז אפן-סאורס כאפט אן Google און OpenAI. אן ענטפער טענהט אז OpenAI'ס ווערע איז דאך שטארק.
- **5 מיי:** **Google, Khan, Slack** - Google האט אויפגעהערט צו פארעפנטליכן פארשונג פאר די פראדוקטן קומען ארויס. Khanmigo, Bard אויף SlackGPT, Pixel.
- **10 מיי:** **IBM, Wendy's** - IBM Watsonx. Wendy's מיט אן AI דרייוו-טרו. Spotify האט אראפגענומען 7% פון Boomy'ס AI-לידער.
- **12 מיי:** **Vela און Meta** - Vela AI-VC מיט 25 מיליאן. Meta AI Sandbox פאר אדווערטייזער.
- **16 מיי:** **Together** - 20 מיליאן סיד-רונדע.

- **16 מיי:** Amazon זוכן – ChatGPT-סטיל זוכן.
- **17 מיי:** Zoom – Zoom אינוועסטירט אין Anthropic.
- **18 מיי:** VC און Google – 187.7 מיליאן אין דיפּעיק-סטארטאפּס. Google AI פאר אדווערטייזינג און מעדיצין.
- **19 מיי:** Musk און Meta – Musk האט אינוועסטירט 50 מיליאן, נישט Meta. 100 גיט אוועק איר טעכנאלאגיע.
- **23 מיי:** Apple, Palantir – 30% פון Apple פון AI אפליקאציעס. Palantir זעט רעקארד-פאדערונג.
- **24 מיי:** Anthropic – 450 מיליאן סעריע C.
- **24 מיי:** Product Studio – Google Product Studio. BuzzFeed Botatouille.
- **25 מיי:** Musk, Snowflake – Musk וויל א קאנקורענט. Snowflake קויפט Neeva.
- **26 מיי:** Nvidia – אקציע +28%.
- **29 מיי:** TinyCorp – 5 מיליאן פאר AMD.
- **30 מיי:** Nvidia – Nvidia האט פארגעשטעלט א נייעם סופער-קאמפיוטער און נייע פראדוקטן. דער DGX פארבינדט 256 Grace CPU מיט 256 Hopper GPU, אינאיינעם 100 טעראבייט זכרון. WPP ארבעט מיט Nvidia אויף אדווערטייזינג.
- **31 מיי:** VC – Salesforce און Workday אינוועסטירן אין AI סטארטאפּס. Duolingo זאגט אז AI איז א פריינד.

רעגירונגען, געזעצן און זיכערהייט

- **1 מיי:** GPT4Free – OpenAI סטראשעט מיט א קלאגע.
- **1 מיי:** פאטענטן, איטאליע – דער סופרים קאורט וויל נישט הערן דעם פאל פון AI ווי ממציא. איטאליע האט צוריק ערלויבט ChatGPT.
- **2 מיי:** כינע – AI מוז ווייזן "סאציאליסטישע ווערטן". כינע איז פאר אמעריקע אין רעגולירונג.
- **3 מיי:** קאפירעכט – דער Warhol פאל, און UMG בעט צו בלאקירן.
- **5 מיי:** ווייסע הויז – מאראלישע פליכט פון די CEO'ס. שטודיע פון אויפזיכט אויף ארבעטער.
- **8 מיי:** GDPR – OpenAI'ס צרות.
- **9 מיי:** OpenAI כללים – די פענטאגאן. OpenAI טריינירט נישט אויף API-דאטא. פארבאט אויף פאליטיק.
- **11 מיי:** Altman פארן קאנגרעס – Altman האט געבעטן רעגולירונג און גערופן AI א "דרוק-מאשין מאמענט".
- **12 מיי:** EU – דער AI Act איז א שטאפל נענטער. Microsoft עקאנאמיסט Schwarz זאגט צו ווארטן ביז "באדייטנדער שאדן".

- **16 מיי:** דיפפעיִקס - Google קעמפט קעגן פאלשע אינהאלט. OpenAI רעד-טיים. Microsoft וואסער-צייכן.
- **19 מיי:** Bengio - פאדערט רעגולירונג און באשרייבט ווי א "ראג" AI קען אויפקומען.
- **19 מיי:** איטאליע - 30 מיליאן יורא פאר ארבעטלאזע.
- **22 מיי:** NYC און G7 - ניו יארק שולן האבן אראפגענומען דעם ChatGPT פארבאט. G7 רופט צו אינטערנאציאנאלע סטאנדארטן.
- **23 מיי:** סופער-אינטעליגענץ - OpenAI זאגט פאראויס אז AI קען אין א צענדליק יאר איבערשטייגן מומחים, און שלאגט פאר אן אינטערנאציאנאלע אויטאריטעט.
- **23 מיי:** פענטאגאן שווינדל - א פאלשער ווידעא פון אן עקספלאזיע ביים פענטאגאן.
- **24 מיי:** Pichai - אן ארטיקל איבער פאראנטווארטליכקייט.
- **26 מיי:** Microsoft - Brad Smith רופט צו א נייע אגענטור און לייסענסעס.
- **26 מיי:** OpenAI אין אייראפע - Altman האט געסטראשעט צו פארלאזן אייראפע, און דערנאך זיך צוריקגעצויגן. ער איז אפטימיסטיש.
- **29 מיי:** אדוואקאט - א לאיער האט ציטירט פאלשע פעלער וואס ChatGPT האט אויסגעטראכט.
- **31 מיי:** סכנה פון אויסשטארבן - פארשער זאגן אז די סכנה דארף באהאנדלט ווערן ווי מגיפות און נוקלעארע מלחמה. Stability AI שרייבט א בריוו צום סענאט לטובת אפן-סאורס.

אנאליזן און מיינונגען

- **1 מיי:** קאוד און באשלוסן - א פארגלייך פון בעטס אין קאוד. ChatGPT איז כמעט 4 מאל מער קאנסיסטענט ווי מענטשן ביי ביזנעס באשלוסן.
- **2 מיי:** Hinton - Hinton האט פארלאזט Google כדי צו רעדן פריי וועגן סכנות. ער זאגט אז Google האט זיך פאראנטווארטליך געפירט.
- **2 מיי:** Time און MIT - Time ווארנט. MIT Technology Review זאגט אז די צוקונפט איז ספעציאליזירט.
- **3 מיי:** טשעטבאטס און דזשאבס - Wattenberger קריטיקירט טשעטבאטס. מורא פאר דזשאבס.
- **4 מיי:** מוזיק און געזונט - AI - So-Vits-SVC, טעראפיע.
- **8 מיי:** ארבעטער - א פירמע האט נישט אפגעזאגט מענטשן. קאנטראקטארן באקומען 15 דאלאר א שעה. מען רעדט וועגן קארפאראציעס. א Hacker News שמועס.
- **9 מיי:** UBI - UBI, Watson, חינוך.
- **11 מיי:** יחידים, אינדיע - יחידים, אינדישע בעטס.
- **12 מיי:** McKinsey - Ted Chong.
- **15 מיי:** Prompt Injection - Simon Willison, נעגאציע, Pichai.
- **16 מיי:** אן צענזור - Eric Hartford.

- **17 מיי:** מאביל – מאביל, CNET יוניאן, Bard JSON.
- **22 מיי:** דאטא – סיילאוס, 60% Abraham Thomas צופרידן, מעדיצין.
- **23 מיי:** ארבעט – Flux.ai, א רעדאקטאר.
- **25 מיי:** נוירא-נעצן – ענליך צו מענטשן. מאדעראציע.
- **29 מיי:** פראדוקציע – Honeycomb, Goertzel.
- **30 מיי:** פראדוקטן – Reforge, CJR.

ווייטערע ידיעות

- **1 מיי:** קלאן – אן AI קלאן האט אפגענארט א באנק און די פאמיליע.
- **3 מיי:** מעדיצין – ברוכן אין CT, mRNA.
- **9 מיי:** Chai – מיליאן-דאלאר טשעלענדזש.
- **19 מיי:** Bilderberg – AI אויפן סדר היום.
- **23 מיי:** פארפלייצונגען – Google פראגנאזן.
- **26 מיי:** Tako – TikTok.
- **29 מיי:** אנטיביאטיק – abaucin, אויסגעקליבן פון 6,000 קאמפאונדס אין 90 מינוט, ערווארטעט ארום 2030.
- **30 מיי:** Barkour – Google ראבאט-טעסט.
- **30 מיי:** מוח און מוזיק – א פאראליזירטער שרייבט 66 אותיות א מינוט. UMG ארבעט מיט Endel.

אין יוני 2023 האט OpenAI ארויסגעלאזט function calling, ביליגערע פרייזן און וועב-זוכן אין ChatGPT. Sam Altman האט געפירט א לאביאיסטישע קאמפיין אין וואשינגטאן, און OpenAI האט געלאביאירט אין איראפע. די אפן-סאורס וועלט איז שטארק געוואקסן: Falcon איז געווארן פריי פאר קאמערציעלן באנוץ, און עס זענען ארויס MPT-30B, OpenLlama, RedPajama און Meta's מוזיק-און קול-מאדעלן. די גרויסע געלט-ידיעות זענען Inflection's 1.3 ביליאן דאלאר, Databricks's איינקויפן פון MosaicML פאר 1.3 ביליאן, Cohere's 270 מיליאן און Mistral's 113 מיליאן. רעגירונגען האבן זיך גענומען צו רעגולירן: דאס איראפעאישע פארלאמענט האט דורכגעלאזט א געזעץ-פלאן, Schumer האט פארגעשטעלט א ראמען, און אין קאנגרעס זענען אריינגעברענגט געווארן נייע ביל'ס.

מאדעלן און נייע לאנסירונגען

- **1 יוני: Altman's פלאן פאר OpenAI – Sam Altman** האט באשריבן OpenAI's פלענער. פאר 2023: א ביליגערע און שנעלערע GPT-4, לענגערע קאנטעקסט-פענצטער, א fine-tuning API און a stateful API. אין 2024 וועט קומען מולטימאדאליטעט.
- **1 יוני: Falcon ווערט פריי** – די Falcon מאדעלן, וואס שטייען #1 אויף דער Hugging Face טאבעלע, זענען פריער ארויסגעלאזט געווארן מיט א באגרעניצנדע הכנסה-טיילונג. יעצט איז די באגרעניצונג אוועקגענומען, און Falcon 40B און 7B זענען אונטער Apache-2, פריי אויך פאר קאמערציעלן באנוץ. Hugging Face האט Falcon אריינגעשטעלט אין איר סיסטעם, צו טריינירן און לויפן אויף CPUs, GPUs און TPUs.



Falcon ווערט פריי · 1 יוני 2023

- **6 יוני: Google's AI אין Gmail און זוכן** – Google האט אנטוויקלט א נייע AI-פונקציע צו זוכן אימעילס שנעלער אין Gmail אויף סמארטפאונס. די נייע AI-זוך, Search Generative Experience, איז געווענליך נוצלעך, אבער די רעזולטאטן נעמען צו לאנג צו לאדן.

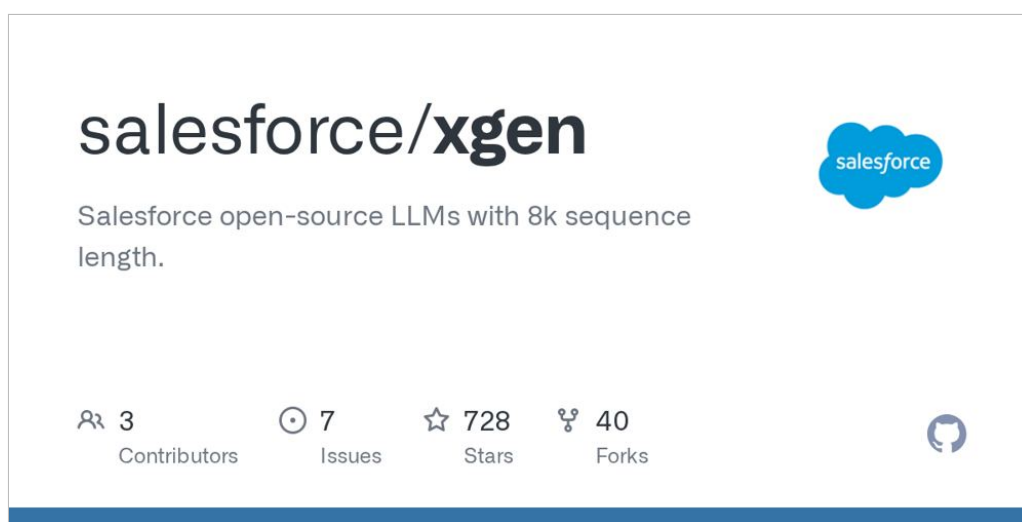
- **6 יוני: Apple's פארשטעקטע AI** – Apple האט ביי WWDC גאר נישט געזאגט "AI", אבער אירע נייע פראדוקטן זענען דורכגעווייקט דערמיט. למשל, די autocorrect אין iOS 17 ניצט א טראנספארמער שפראך-מאדעל. Ars Technica שרייבט אז Apple קען זיין דער שטילער פירער אין AI פאר געווענליכע קאנסומערס.
- **7 יוני: RedPajama 7B פארטיג** – דער RedPajama 7B מאדעל האט פארענדיגט זיין טריינירן אויף 3000 V100s. די instruction-tuned ווערסיע איז דער שטארקסטער אפענער מאדעל פון איר גרייס.
- **12 יוני: Meta's MusicGen** – Meta האט ארויסגעלאזט MusicGen, א מאדעל וואס מאכט מוזיק פון טעקסט אדער פון א מעלאדיע. דער קאוד איז אינגאנצן אפן-סאורס, אבער די מאדעל-וואגן מעג מען נישט ניצן קאמערציעל. MusicGen איז א טראנספארמער מיט איין שטאפל, וואס זאגט פאראויס אלע codebooks גלייכצייטיג, און דאס פארמינערט די צאל טריט פאר יעדע סעקונדע קלאנג. ער געפינט זיך אין דער Audiocraft ביבליאטעק.
- **12 יוני: Khanmigo** – Khan Academy האט לאנסירט Khanmigo, א טשעטבאט אויף GPT-4 וואס לערנט מיט סטודענטן פערזענלעך מיט דער סאקראטישער מעטאד. מען זאגט זיך אבער וועגן טעותים און וועגן דער השפעה אויף לערער.
- **13 יוני: WordPress Jetpack AI** – WordPress האט ארויסגעלאזט Jetpack AI אויף GPT-3, וואס שרייבט בלאג-פאסטס און סאשעל-מידיע פאסטס. דאס געצייג איז נאך אין בעטא.
- **13 יוני: Bard לויפט קאוד** – Bard וועט באלד קענען עקספארטירן ענטפערס צו Google Sheets. ער וועט אויך אליין דערקענען ווען א פראגע פאדערט מאטעמאטיק אדער לאגיק, און דאן לויפן קאוד אינעווייניג פאר מער פינקטליכע ענטפערס.
- **14 יוני: OpenAI function calling** – OpenAI האט ארויסגעלאזט נייע ווערסיעס פון gpt-3.5-turbo און GPT-4 מיט function calling: א פראגראמירער באשרייבט פונקציעס, און דער מאדעל שאפט קאוד וואס רופט זיי אויף. דער פרייז פון gpt-3.5-turbo איז אראפ מיט 25%. לויט every.to איז דאס א גרויסע זאך, ווייל GPT-4 קען יעצט ניצן דרויסנדיגע געצייג, ווי ארויסגעפינען דאס וועטער, א סטאק-פרייז אדער דאטא פון א פירמע.
- **14 יוני: Meta's I-JEPA** – Meta האט פארגעשטעלט I-JEPA, דעם ערשטן מאדעל לויט Yann LeCun'ס גרויסער וויזיע. דער מאדעל לערנט זיך אן אינערליכן בילד פון דער וועלט דורך פאראויסזאגן שטיקער פון בילדער, ענליך ווי מענטשן לערנען זיך. LeCun איז לאנג שוין סקעפטיש אויף דעם היינטיגן וועג פון AI פארשונג.
- **15 יוני: Google AI ביים איינקויפן** – Google ברענגט גענעראטיווע AI אין אירע שאפינג-פונקציעס, כדי צו פארכאפן דעם מארק פאר Amazon.
- **16 יוני: Mercedes מיט ChatGPT** – Mercedes-Benz שטעלט אריין ChatGPT אין איר MBUX סיסטעם אין קאר, פאר מער נאטירליכע שמועסן מיט די דרייווערס און פאסאזשירן.
- **19 יוני: OpenLlama** – OpenLlama איז א נאכמאך פון Meta's Llama אויף דעם RedPajama 1T דאטא-סעט, מיט א פרייע ליצענץ. די 3B, 7B און 13B מאדעלן ארבעטן כמעט ווי די אריגינעלע.



2023 יוני 19 · OpenLlama

- **19 יוני:** Meta Voicebox – Meta האט פארגעשטעלט Voicebox, דער בעסטער מאדעל היינט פאר טעקסט-צו-רייד.
- **21 יוני:** Opera's AI בלעטערער – Opera זאגט אז איר בלעטערער מיט גענעראטיווע AI איז גרייט פארן ציבור.
- **23 יוני:** Inflection's מאדעל – Inflection האט ארויסגעוויזן דעם שפראך-מאדעל הינטער איר Pi טשעטבאט.
- **23 יוני:** ElevenLabs – ElevenLabs האט אויפגענומען 19 מיליאן דאלאר אין א Series A, געפירט פון Nat Friedman, Daniel Gross און Andreessen Horowitz. זי האט מער ווי 1 מיליאן רעגיסטרירטע באניצער. זי האט אויך לאנסירט Voice Library, א פלאץ צו שאפן און טיילן קינסטלעכע קולער לויט עלטער, אקצענט און אזוי ווייטער.
- **23 יוני:** SDXL 0.9 – Stability AI האט ארויסגעלאזט SDXL 0.9, וואס שאפט מער רעאליסטישע בילדער מיט מער פרטים אין דער קאמפאזיציע. זי האט אויך ארויסגעלאזט צוויי נייע דיפוזיע-מאדעלן פאר פארשונג.
- **23 יוני:** MPT-30B – MosaicML האט ארויסגעלאזט MPT-30B. ער איז טריינירט אויף H100s, האט א קאנטעקסט פון 8k, א קאמערציעלע ליצענץ, קען שרייבן קאוד, און פאסט אויף איין A100. דער 7B מאדעל פון MosaicML איז שוין אראפגעלאדן געווארן מיליאנען מאל.
- **26 יוני:** Google AudioPaLM – AudioPaLM פארבינדט א טעקסט-מאדעל מיט א רייד-מאדעל אין איין מולטימאדאלן מאדעל. ער ווערט גענוצט פאר רייד-דערקענונג און איבערזעצונג פון רייד צו רייד.
- **26 יוני:** Midjourney 5.2 – Midjourney ברענגט א "zoom out" פונקציע: דאס בילד אין צענטער בלייבט, און רונד ארום ווערט אויסגעבויט א גרעסערע סצענע.
- **27 יוני:** AI אין Colab – Google ברענגט איר AI קאוד-הילף אין Colab פאר Pro+ אבאנענטן.
- **28 יוני:** Ernie קעגן ChatGPT – Baidu טענהט אז Ernie Bot איז שטארקער ווי ChatGPT בכלל, און אין עטליכע כינעזישע פעאיגקייטן שלאגט ער אפילו GPT-4.

- **28 יוני: Kosmos-2** – Microsoft's UniLM גרופה האט ארויסגעלאזט Kosmos-2, א מולטימאדאלער מאדעל וואס קען צייכענען bounding boxes און לייענען טעקסט אין בילדער.
- **29 יוני: ChatGPT זוכט אויפן וועב** – ChatGPT Plus אבאנענטן קענען יעצט לאזן ChatGPT זוכן אין Bing, פאר נייעס און אינפארמאציע פון נאך זיין טריינירן. מען הייבט אבער די ברעמען וועגן ניצן נאר Bing, צוליב OpenAI's נאנטע באציאונגען מיט Microsoft.
- **29 יוני: Open Flamingo v2** – LAION האט ארויסגעלאזט Open Flamingo v2, א נאכמאך פון DeepMind's Flamingo אויף MPT און RedPajama. ער איז אומגעפער 80% אזוי שטארק ווי דער אריגינעל, און דער קאוד און די וואגן זענען צוגעגלייך.
- **29 יוני: Salesforce's 7B מאדעל** – Salesforce האט ארויסגעלאזט 7B מאדעלן מיט 4K און 8K קאנטעקסט, טריינירט אויף 1.5 טריליאן טאקעס. זיי זענען אזוי גוט ווי MPT און Falcon אדער בעסער. די instruct ווערסיע איז נאר פאר פארשונג.



Salesforce's 7B מאדעלן · 29 יוני 2023

- **30 יוני: Bing שאפינג** – Microsoft האט ארויסגעלאזט נייע AI שאפינג-געצייג אין Bing און אין Edge.

פארשונג און טעכניק

- **1 יוני: פראצעס-השגחה אין מאטעמאטיק** – OpenAI's "process supervision" טריינירט דעם מאדעל אויף יעדן טריט פונעם חשבון און נישט נאר אויפן ענטפער. דאס ברענגט א שטארקע פארבעסערונג ביי שווערע מאטעמאטישע פראבלעמען.
- **1 יוני: פארגיפטן טריינינג-דאטא** – א פארשונג-ארבעט ווייזט אז מען קען אריינלייגן שעדליכע ביישפילן אין די טריינינג-דאטא, אזוי אז דער מאדעל פירט זיך מאדנע אויף ווען מען דערמאנט געוויסע ווערטער.
- **1 יוני: דיפוזיע שפראך-מאדעלן** – מיט שליווע ענדערונגען (score matching) האבן פארשער באקומען א 1B דיפוזיע שפראך-מאדעל מיט likelihood scaling laws, וואס ארבעט אזוי גוט ווי א געווענליכער מאדעל 10 מאל קלענער.

- **1 יוני:** עטליכע פארשונגען – א מעטאד צו שאפן סטיילישע 3D אוואטארן מיט דיפוזיע און GAN. מאדעלן וואס שאפן אליין זייערע אייגענע געצייג ("utility functions") אנשטאט ניצן פארטיגע. Macaw-LLM, וואס פארבינדט בילד, ווידעא, קלאנג און טעקסט אויף CLIP, Whisper און LLaMA.
- **2 יוני:** Google's קאוד-היסטאריע – Google האט לאנגע יארן רעקארדירט יעדע ענדערונג אין איר קאוד. זי ניצט דאס יעצט אלס sequence modeling, צו בויען אינערליכע געצייג וואס מאכן איר פראגראמירער מער פראדוקטיוו.



Google's קאוד-היסטאריע · 2 יוני 2023

- **2 יוני:** קליינערע און ביליגערע מאדעלן – knowledge distillation איז אפט מער כדאי ווי געלט ארויסגעבן אויף מער דאטא. א צווייטע פארשונג-ארבעט שטודירט וואס עס טוט זיך ווען מען קוואנטיזירט פאר'ן טריינירן, און טאפט ארום די באוואוסטע גרעניץ פון 6B פאראמעטערס.
- **2 יוני:** בילד און ווידעא – GeneFace++ שאפט דאס ערשטע מאל רעדנדע 3D פנימער אין רעאלער צייט. Occ-BEV ניצט כמה קאמערעס פאר אויטאנאמישע קארן. GPT4Tools שטייערט כמה וויזועלע מאדעלן מיט LoRA.
- **5 יוני:** Brainformer – Google האט געזוכט מיט א גענעטישן אלגאריטם אויף א הויפן TPU, און געפונען א מאדעל מיט MoE בלאקן וואס לערנט זיך 5 מאל שנעלער און לויפט 2 מאל שנעלער ווי א טראנספארמער.
- **5 יוני:** בילד-שאפונג – ArtGPT-4 פארשטייט און שאפט קונסט-בילדער, און קומט נאנט צו קינסטלער אין געוויסע טעסטן. BlendFields מאכט פיינערע פנים-פרטים. StyleDrop שאפט בילדער אין יעדן סטיל פון איין ביישפיל-בילד.
- **6 יוני:** נאך פארשונגען – ViCo – בילד-שאפונג פון טעקסט אן GateON. fine-tuning – לערנען נייע אויפגאבעס אן פארגעסן די אלטע. Prompt-Free Diffusion – בילדער נאר פון וויזועלע אינפוטס, אן טעקסט.

- **7 יוני: דאטא און וויזיע** – StableRep ניצט קינסטלעכע בילדער פון Stable Diffusion און שלאגט SimCLR און CLIP. א נייער קארפוס פון כמעט 100 ביליאן טאקענס פון געזעץ-דאקומענטן אין 24 שפראכן (689GB) ברענגט נייע רעקארדן. MotionDiffuser זאגט פאראויס באוועגונגען פון כמה קארן אדער ראבאטן.
- **8 יוני: AlphaDev און DeepMind** – DeepMind's AlphaDev באהאנדלט סארטירן ווי א שפיל פאר איין שפילער, און האט געפונען א שנעלערן סארטיר-אלגאריטם. AlphaZero האט פארבעסערט די צוטיילונג פון ארבעט אין דאטא צענטערס, אזוי אז ווייניגער הארדווער שטייט ליידג (19%). MuZero האט פארבעסערט YouTube's ווידעא-קאמפרעסיע, און AlphaDev's hashing האט אראפגעברענגט די לאסט מיט 30%.
- **8 יוני: בילד-פארשונג** – Stable Diffusion לערנט זיך אליין צו דערקענען וועלכע טיילן פון פארשידענע בילדער זענען דאס זעלבע. CiT-Net – מעדיצינישע בילד-סעגמענטאציע. א פארשונג געפינט אז סופערווייזד וויסן קען אמאל שאטן ביי Hierarchical Novel Class Discovery. vision – א פשוטערער vision transformer, שנעלער און מער פינקטלעך.
- **9 יוני: ChatGPT און וויצן** – דייטשע פארשער האבן געפונען אז 90% פון 1,008 וויצן וואס ChatGPT-3.5 האט געשאפן זענען געווען די זעלבע 25 וויצן. ער קען ערקלערן וויצן, אבער ער האט שוועריגקייטן צו שאפן נייע.
- **9 יוני: עטליכע ארבעטן** – ערשטער פלאץ אין Git-Theta, video panoptic segmentation, אן הרחבה פון Git פאר ML מאדעלן. Youku-mPLUG, דער גרעסטער עפנטליכער כינעזישער ווידעא-טעקסט דאטא-סעט, מיט 10 מיליאן פארן.
- **12 יוני: מולטימאדאל, קאמפרעסיע, אמת** – Otter פון סינגאפור איז א נייער מאדעל פאר טעקסט און בילד. SpQR קאמפרעסירט LLMs כמעט אן פארלוסט, מער ווי 4 מאל, אזוי אז זיי קענען לויפן אויף לעפטאפס און טעלעפאנען. דורך ענדערן די אקטיוואציעס האט Alpaca's TruthfulQA רעזולטאט זיך געהויבן פון 32.5% ביז 65.1%.
- **12 יוני: וויזיע** – Matte Anything ניצט דריי בילד-מאדעלן פאר image matting. OmniMotion נאכפאלגט יעדן פיקסל אין א ווידעא אויף לאנגע צייט, אפילו ווען ער ווערט צוגעדעקט.
- **13 יוני: מאדעלן קריטיקירן** – CID ווייזט ווי מאדעלן רעאגירן אויף קליינע ענדערונגען אין אינפוט. אין RL4F העלפט א קליינער קריטיק-מאדעל דעם גרויסן GPT-3 פאררעכטן טעותים, מיט א פארבעסערונג פון 5% אין דורכשניט. Hugging Face ווייזט אז GPT-4 אלס ריכטער האט ליב לענגערע ענטפערס, און דאס ברענגט אריין פראגעס וועגן די פארגלייכן מיט ChatGPT.
- **14 יוני: GAP און AlgoPerf** – א פאר טריט Gradient Ascent אויף צופעליגן טעקסט פארבעסערט קליינע מאדעלן ביז זיי ארבעטן ווי גרעסערע. AlgoPerf איז א בענטשמארק צו פארגלייכן טריינינג-אלגאריטמען.
- **15 יוני: עפיציעניץ** – SoViT, א vision transformer מיט אן אפטימירטער גרייס, איז קאנקורענטפלאג מיט פיל גרעסערע מאדעלן. ווייטער: א בעסערער RGB-D דעטעקטאר, און CAN פאר דערקענען אנאמאליעס אין time series.

- **15 יוני: מאדעל-צוזאמענפאל** – פארשער ווארנען אז מאדעלן טריינירט אויף AI-געשאפענעם אינהאלט "פאלן צוזאם": זיי פארגעסן די עכטע דאטא און ווערן פולער טעותים. דאס איז אבער באוויזן געווארן נאר אויף א מאדעל פון 125M פאראמעטערס.
- **16 יוני: Waymo און LoRA** – Waymo ניצט דיפוזיע פאר פאראויסזאגן באוועגונגען אויפן וועג, בעסער ווי פיזיקאלישע מעטאדן. א נייער מאדולארער שיכט מאכט fine-tuning שנעלער אויף ביליגער הארדווער.
- **19 יוני: נאך ארבעטן** – ווי אזוי מען קען צוזאמשטעלן א היימישן A100 סופער-קאמפיוטער פאר עטליכע טויזנט דאלאר. גרויסגעמאכטע value-based RL דערגרייכט העכער ווי מענטשן. א גרויסער מאדעל קען לערנען א קלענערן ("לערער-תלמיד").
- **19 יוני: 100K קאנטעקסט** – די טעכניקן הינטער 100K טאקענס קאנטעקסט זענען ALiBi, Sparse Attention, FlashAttention, Multi-Query attention, conditional computation, און 80GB A100s.
- **20 יוני: עטליכע ארבעטן** – FALL-E שאפט קלאנג פון טעקסט. MAGICBRUSH איז א דאטא-סעט פון מער ווי 10,000 ביישפילן פאר בילד-רעדאגירן לויט אנווייזונגען. א multi-task טראנספארמער פאר פארשטיין סצענעס. א טראנספארמער מיט וואגן אויסגעקליבן מיט דער האנט איז 100% פינקטלעך אין לאנגע חשבונות פון צולייגן.
- **21 יוני: DeepMind RoboCat** – געבויט אויף Gato, איז א ראבאט-אגענט וואס לערנט זיך א נייע אויפגאבע אין ווייניג ווי 100 טריט.
- **21 יוני: נאך ארבעטן** – PhotoBook – מאשינען לערנען זיך שמועסן ווי מענטשן וועגן בילדער. LOMO – fine-tuning פון א 65B מאדעל אויף איין מאשין מיט RTX 3090 8 טשיפס, מיט נאר 10.8% פון דעם זכרון. Clip2Protect – נאטירלעך-אויסזעענדע "מעיקאפ" וואס נארט אפ פנים-דערקענונג.
- **22 יוני: "טעקסטבוקס" פאר קאוד** – א מאדעל פון 1.3B, טריינירט אויף גוט-אויסגעקליבענע דאטא און קינסטלעכע לערנביכער מיט איבונגען, באקומט מער ווי 50 אויף HumanEval. דאס איז נאנט צו ChatGPT און פיל בעסער ווי דער Codex 12B.
- **22 יוני: האלוצינאציעס און עפיציענץ** – SequenceMatch באהאנדלט שרייבן טעקסט ווי RL און פארמינערט האלוצינאציעס. א נייע ארדענונג פון זכרון אויף GPU מאכט טראנספארמערס פיל מאל שנעלער. Wanda שניידט אפ אומוויכטיגע וואגן אן נאכאמאל טריינירן.
- **23 יוני: נאך ארבעטן** – Tart, א מאדול וואס פארבעסערט פראבאביליסטישע שכל-ארבעט ביי יעדן מאדעל. א קליינער ResNet מיט ווייניגער ווי 5 מיליאן פאראמעטערס דערגרייכט 96.04% אויף CIFAR-10. אן אלטערנאטיוו צו SAM, 50 מאל שנעלער.
- **26 יוני: ראבאטן און קאוד** – SoftGPT לערנט ראבאטן אנפאסן ווייכע זאכן ווי בגדים. disentanglement קומט צוריק פאר רעדאגירן בילדער מיט דיפוזיע. א פארשונג געפינט אז GPT-4 איז כמעט דער איינציגער מאדעל וואס קען פאררעכטן זיין אייגענעם קאוד. דער שרייבער פון ניוולעטער שאצט אז OpenAI האט אים ספעציעל אזוי טריינירט.
- **27 יוני: בילדער און שטעט** – PromptIR שטעלט צוריק בילדער פון יעדן סארט שאדן (ראש, בלור, נעפל). SHGNN מאכט שטאט-מאפעס וואס האלטן אכט אויף דעם חילוק צווישן ערטער.

- **28 יוני:** אפשאצונג און דאטא-סעטן – self-supervised evaluation לאזט אפשאצן מאדעלן אויף די אייגענע דאטא, אפילו אויף דאטא וואס קומט אריין פון פראדוקציע. LRV-Instruction מינערט האלוצינאציעס ביי מולטימאדאלע מאדעלן. FunQA, א דאטא-סעט פון חידושיגע ווידעאס.
- **29 יוני:** נאך ארבעטן – 3D-Speaker – קלאנג-דאטא פון מער ווי 10,000 רעדנער. א פארשונג וועגן טריינירן אויף פשוטערער שפראך. LightGlue – פארגלייכט בילדער שנעלער ווי SuperGlue.
- **30 יוני:** נאך ארבעטן – ChatLaw, אן אפן-סאורס מאדעל פאר כינעזישע געזעצן. אן ארטיקל וועגן "integration" קעגן "coverage": קליינע מאדעלן קענען זיין גוט אויף בענטשמארקס דורך מער אויסנווייניג-וויסן. א איבערבליק פון instruction tuning, צוזאמען מיט דער ערקלערונג פארוואס מען שאצט FLAN נישט גענוג. XrayGPT לייענט ברוסט-רענטגענס. Lens גיט א פארפרוירענעם שפראך-מאדעל צו פארשטיין בילדער דורך קליינע מאדולן.

געצייג פאר פראגראמירער און אפן-סאורס

- **1 יוני:** געצייג – א קורס פאר LangChain. Siit ענטפערט אויף פראגעס פון ארבעטער פון Notion און Confluence, אויף GPT-4, Recraft, פריי וועקטאר- און 3D-קונסט, מיט ערלויבעניש פאר קאמערציעלן באנוץ.
- **2 יוני:** געצייג – SmartGPT און SuperAGI פאר AI אגענטן. א Key/Value cache מאכט MosaicML's MPT Storywriter מיט 65k קאנטעקסט כמעט 10 מאל שנעלער. Speechki גיט ChatGPT א קול.
- **5 יוני:** געצייג – AWQ, קוואנטיזאציע וואס היט אפ די וויכטיגסטע וואגן. CodeTF, א Python ביבליאטעק פאר קאוד-מאדעלן. AITemplate פארוואנדלט נעצן אין C++/HIP/CUDA. Poe באט פאר Midjourney פראמפטס. Zeda.
- **6 יוני:** געצייג – Tree of Thoughts, וואס פארבעסערט שכל-ארבעט מיט 70% לויט זיינע שאפער. OpenAI QLoRA's עצות ווי צו ניצן GPT. Market Test.
- **7 יוני:** געצייג – Aviary פארגלייכט LLMs, צוזאמען מיט קאסטן און שנעלקייט. InstructZero פארבעסערט אנווייזונגען פאר "black-box" מאדעלן. Sivi AI. Blobr. HQ-SAM. מאכט פון יעדן API א ChatGPT פלאגאין אין 2 מינוט.
- **8 יוני:** געצייג – uBlock כלל וואס ווייזט נאר טוועעטס מיט Arxiv לינקס. פרייע קורסן פון DeepLearning.AI. ClientZen. Pitch Avatar. ChatHub.
- **9 יוני:** געצייג – Bard API פאר Python. Simon Willison's GPT tokenizers. TeamGPT. Supabase Vector. HearHear.
- **12 יוני:** געצייג – Lanarky. ClientZen. Gong Engage.
- **13 יוני:** געצייג – DataDM, א לאקאלער דאטא-געהילף. FinGPT, אן אפן-סאורס מאדעל פאר פינאנצן. א איבערבליק וועגן LLMs אין רעקאמענדאציע-סיסטעמען. Composer פאר האנדל-אלגאריטמען. LocalAI. ChatALL.
- **14 יוני:** געצייג – א LLaMA tokenizer אין JavaScript. detrex. ChatCSV. Assembly.

- **15 יוני: געצייג** – OmniEvent. WizardCoder, אן אפענער מאדעל אויף Starcoder מיט Evol אנווייזונגען, וואס דערגרייכט 57 אויף HumanEval און שלאגט ChatGPT ביי קאוד. Framer AI.
- **16 יוני: געצייג** – Easy OCR פאר מער ווי 80 שפראכן. GPT. Hindsight Experience Replay. Engineer, וואס שאפט גאנצע קאוד-באזעס פון א פראמפט. AlphaCTR. Lance. BeforeSunset.
- **19 יוני: געצייג** – Seal פאר קאר-סענסארן, מיט כמעט 37% פארבעסערונג. Python בויסטיינער וואס לויפן אויף Stanford. Jax ראמען פאר foundation מאדעלן. AI QR קאודס מיט ControlNet. Prompt Flat.
- **20 יוני: געצייג** – simpleaichat. ML. SSCBench. פאר האנדל: 800 זייטן און 150 נאטבוקס. Vercel AI. SDK. Olvy 2.0.
- **21 יוני: געצייג** – Autolabel. ווי אזוי צו דעוּפליקירן גרויסע דאטא-סעטן. EnCodec איז יעצט אין Hugging Face transformers. Typefully 2.0. Quivr.
- **22 יוני: געצייג** – ClickUp AI. ScholarTurbo. OpenSTL.
- **23 יוני: געצייג** – Kaguya, א ChatGPT פלאגאין פאר לאקאלע פיילס און סקריפטס. FlagAI. Upword. 2.0.
- **26 יוני: געצייג** – Microsoft's DeepSpeed האט באקומען א פארבעסערונג אין זכרון, שנעלקייט און באקוועמקייט. 3HAN דערקענט פאלשע נייעס. Codium. Boring Report. OpenMMLab. ExLlama.
- **27 יוני: געצייג** – LMFlow, א ביבליאטעק פאר fine-tuning. דער קאוד פון DragGAN איז ארויס. Hugging Face. Giskard's QR קונסט-געצייג.
- **28 יוני: געצייג** – FiftyOne. MotionGPT. אן איבערבליק וועגן AI אגענטן אויף LLMs.
- **29 יוני: געצייג** – MosaicML's Goodmeetings. Composer.
- **30 יוני: געצייג** – Talk-Codebase. א קעיס-סטאדי פון בויען Casemark. Virtual CMO. Boba AI. פאר אדוואקאטן.

פירמעס, געלט און ביזנעס

- **2 יוני: Baidu's פאנד** – Baidu עפנט א פאנד פון 1 ביליאן יואן (145 מיליאן דאלאר) פאר AI סטארטאפס, צוזאמען מיט א קאנקורס פאר אפליקאציעס אויף ERNIE.
- **2 יוני: Amazon** – Amazon פירט איין AI אין א טויז לאגערן, צו כאפן געשעדיגטע סחורה איידער מען שיקט זי ארויס.
- **5 יוני: Asana און Apple** – Asana Intelligence ברענגט AI פאר רעסורסן-פירונג, אויטאמאטישע workflows און קורצע איבערבליקן. Apple זוכט אינוועסטירן פאר גענעראטיווע AI.
- **6 יוני: OpenAI און דער AI בום** – נאך ChatGPT האבן ווענטשער קאפיטאליסטן אינוועסטירט מער ווי 11 ביליאן דאלאר אין AI סטארטאפס. לויט א פאסט, וואס איז שפעטער אראפגענומען געווארן, האט Altman פריוואט צוגעזאגט פראגראמירער אז OpenAI וועט נישט ארויסלאזן נאך קאנסומער-פראדוקטן חוץ ChatGPT.

- **7 יוני: פירמעס** – Altman וויל נישט נעמען OpenAI עפנטלעך, כדי צו האלטן קאנטראל. GGML, וואס לויפט מאדעלן אויף MacBooks, איז געווארן א פירמע. Instagram ארבעט אויף א טשעטבאט.
- **9 יוני: Cohere און Meta** – Cohere האט אויפגענומען 270 מיליאן דאלאר אין א Series C. Mark Zuckerberg האט אנגעזאגט פאר זיינע ארבעטער אז Meta וועט אריינשטעלן גענעראטיווע AI פאר טעקסט, בילדער און ווידעא אין Facebook און Instagram.
- **9 יוני: Adobe פארזיכערט** – Adobe וועט באצאלן פאר ביזנעסער אויב מען קלאגט זיי פאר copyright צוליב בילדער פון Firefly.
- **13 יוני: LlamaIndex און Salesforce** – LlamaIndex האט אויפגענומען 8.5 מיליאן דאלאר seed, געפירט פון Salesforce. Greylock טאפלט איר אינוועסטירונג אין AI סטארטאפס.
- **14 יוני: Mistral AI** – Mistral AI, א סטארטאפ אין פאריז נאר פיר וואכן אלט, געגרינדעט פון געוועזענע ארבעטער פון DeepMind און Meta, האט אויפגענומען 113 מיליאן דאלאר seed. די רונדע איז געפירט געווארן פון Lightspeed, ביי א ווערט פון 260 מיליאן. זי צילט צו אפן-סאורס מאדעלן פאר פירמעס.
- **14 יוני: Microsoft און OpenAI** – הינטער די קוליסן ברענגט די שותפות צווישן Microsoft און OpenAI אויך קאנפליקטן. Satya Nadella שטעלט אלעס אויף AI, און די 10 מיליאן דאלאר אינוועסטירונג אין OpenAI וועט ווערן זייער פראפיטאבל. ווייטער: אנדזשעל אינוועסטארן האבן אויפגעשטעלט א קלאסטער פון 2,000 H100s פאר סטארטאפס.
- **15 יוני: Synthesia און AMD** – Synthesia האט אויפגענומען 90 מיליאן דאלאר Series C, געפירט פון AMD. Accel וועט שפעטער היי יאר ארויסשיקן דעם MI300X טשיפ פאר LLMs, וואס שטעלט זיך קעגן Nvidia. לויט די פראגנאזן וועט דער מארק וואקסן פון 30 ביליאן ביז מער ווי 150 ביליאן דאלאר ביז 2027.
- **19 יוני: Meta און ByteDance** – Meta פלאנט צו געבן פרייע מאדעלן פאר קאמערציעלן באנוץ, וואס וועט שטעלן דרוק אויף Google און OpenAI. ByteDance האט באשטעלט כמעט 100 טויזנט GPUs פון Nvidia פאר 1 ביליאן דאלאר, נאך פאר די עקספארט-באגרעניצונגען.
- **20 יוני: Snap** – Snap פרובירט צו ניצן איר טשעטבאט פאר אדווערטייזינג.
- **21 יוני: OpenAI אפליקאציע-סטאר** – OpenAI פלאנט א מארק וואו פראגראמירער וועלן קענען פארקויפן מאדעלן געבויט אויף איר טעכנאלאגיע.
- **22 יוני: טשיפס און פירמעס** – Nvidia's A100 און H100 ווערן פארקויפט אויפן אונטערערדישן מארק אין כינע פאר טאפלטן פרייז. Freshworks האט ארויסגעלאזט נייע Freddy AI פונקציעס. Ant Group בויט א LLM.
- **26 יוני: AWS** – AWS עפנט א Generative AI Innovation Center מיט 100 מיליאן דאלאר.
- **27 יוני: Databricks קויפט MosaicML** – Databricks קויפט איין MosaicML, א פירמע פון צוויי יאר, פאר 1.3 ביליאן דאלאר. דאס פארבינדט MosaicML'ס געצייג פאר טריינירן מיט Databricks'ס קונה-דאטא.
- **5 יוני: Stability AI** – א טיפע אויספארשונג אין Forbes וועגן Emad Mostaque'ס פריערדיגע פארגרעסערטע טענות.

- **27 יוני: Stability AI'ס צרות** – דער ראש פון פארשונג David Ha און דער COO Ren Ito האבן פארלאזט Stability AI. די פירמע האט נישט געקענט אויפנעמען געלט ביי א ווערט פון 4 ביליאן, און זי האט באקומען א convertible note פון 25 מיליאן דאלאר.
- **27 יוני: OpenAI אסיסטענט** – OpenAI וויל פארקויפן א "פערזענליכן אסיסטענט פאר ארבעט", וואס וועט זי שטעלן אין קאנקורענץ מיט Microsoft.
- **28 יוני: Reka** – Reka, פון דעם שאפער פון FlanT5, האט אויפגענומען 58 מיליאן דאלאר. זי איז א פירמע וואס ארבעט בעיקר פון דער ווייטנס.
- **28 יוני: Inflection** – Inflection'ס קלאסטער מיט Nvidia און CoreWeave, פון מער ווי 3,000 H100s, האט טריינירט א מאדעל אין GPT-3 סטיל אין 11 מינוט; דער פריערדיגער רעקארד איז געווען 23 טעג. די פירמע, א יאר אלט, האט אויפגענומען 1.3 ביליאן דאלאר, מערסטנס פאר קאמפיוט, און זי האט אנגעזאגט 22 טויזנט H100s.
- **28 יוני: אינוועסטארן** – אינוועסטארן לויפן קויפן שערס פון AI סטארטאפס אויף פריוואטע מערק. די ווענטשער-אפטיילונגען פון גרויסע פירמעס עפענען נייע פאנדס.
- **29 יוני: OpenAI אין לאנדאן** – OpenAI עפנט איר ערשטן אויסלענדישן אפיס אין לאנדאן.

רעגירונגען, געזעצן און זיכערהייט

- **1 יוני: צעטיילט אין ווייסן הויז** – די Biden רעגירונג איז צעטיילט: אין ווייסן הויז און אין Commerce ווילן געוויסע באאמטע א שטרענגע רעגולירונג ווי אין אייראפע, בשעת די מענטשן פון נאציאנאלע זיכערהייט ווילן ווייניגער רעגולירונג, כדי אמעריקע זאל בלייבן קאנקורענטפֿעאיג.
- **2 יוני: OpenAI זיכערהייט-גרענט** – OpenAI עפנט א פראגראם פון 1 מיליאן דאלאר פאר סייער-זיכערהייט פראיעקטן מיט AI.
- **6 יוני: Altman אין וואשינגטאן** – Altman האט זיך באגעגנט מיט מער ווי 100 קאנגרעסלייט, און ער זאגט אז זיין מוט איז געשטארקט פון דעם רצון פון וועלט-פירער צו באגרעניצן די סכנות. OpenAI טריינירט נאך נישט GPT-5, און זי איז "זיכער נישט נאנט" דערצו.
- **8 יוני: גנבה פון GPT-4** – מענטשן כאפן API שליסלען וואס פראגראמירער האבן געלאזט אפן, און ניצן GPT-4 אן צו באצאלן.
- **12 יוני: ביל'ס אין קאנגרעס** – צוויי ביל'ס פון ביידע פארטייען: איינס פאדערט אז די רעגירונג זאל זיין דורכזיכטיג ווען זי ניצט AI, און דאס אנדערע שטעלט אויף אן אפיס פאר קאנקורענץ אין טעכנאלאגיע. Hawley און Blumenthal האבן אריינגעברענגט א ביל אז Section 230 זאל נישט שיצן גענעראטיווע AI.
- **13 יוני: פאראויסדיגער צוטריט פאר UK** – OpenAI, Google DeepMind און Anthropic האבן צוגעזאגט צו געבן בריטאניע פריערדיגן צוטריט צו זייערע מאדעלן פאר זיכערהייט-פארשונג.
- **14 יוני: Google קעגן נייעם רעגולאטאר** – Google וויל צעטיילן די השגחה איבער AI צווישן פארשידענע אגענטורן, אנשטאט א באזונדערן רעגולאטאר ווי Microsoft און OpenAI ווילן.

- **15 יוני: אייראפעאיש געזעץ** – דאס אייראפעאישע פארלאמענט האט דורכגעלאזט א פלאן פון א געזעץ, וואס באגרעניצט AI אין סכנה-פולע געביטן ווי דאס געריכט-סיסטעם, און פאדערט מער אויפקלערונג וועגן די דאטא.
- **15 יוני: Bard און פריוואטקייט** – Google האט געמוזט אפשטופן Bard אין אייראפע, ווייל די אירישע Data Protection Commission האט נישט באקומען קיין פולע אינפארמאציע וועגן פריוואטקייט. Google האט אויך געווארנט אירע ארבעטער נישט אריינצולייגן קאנפידענציעלע אינפארמאציע אין טשעטבאטס, אריינגערעכנט Bard.
- **19 יוני: כינעזישע AI** – די Biden רעגירונג ראנגלט זיך מיט דער פראגע וועלכע AI איז א סכנה פאר נאציאנאלע זיכערהייט, בשעת זי באגרעניצט אינוועסטירונגען אין כינע.
- **21 יוני: OpenAI לאביאירט אין EU – OpenAI** האט געלאביאירט אז מען זאל אפשרוואכן טיילן פונעם EU AI Act, כדי צו פארמינערן די לאסט אויף איר.
- **21 יוני: אינוועסטארן דריקן** – גרויסע אינוועסטארן דריקן אויף טעק-פירמעס צו נעמען אחריות פאר מיסברויך פון AI.
- **22 יוני: ווייסן הויז און Schumer** – AI איז געווארן א שפיץ-פריאריטעט אין ווייסן הויז, וואו באאמטע טרעפן זיך כמה מאל א וואך וועגן פאלשע אינפארמאציע, סייבער און עקאנאמיע. מען פלאנט עקזעקיוטיוו-באפעלן אין זומער. Chuck Schumer האט פארגעשטעלט זיין ראמען פאר רעגולירן AI.
- **28 יוני: אייראפעאישע "קראש טעסטן"** – די EU עפנט צענטערס צו טעסטן AI פאר זיכערהייט איידער זי קומט אויפן מארק.
- **30 יוני: ChatGPT אין קאנגרעס** – Ted Lieu קריטיקירט די נייע כללים אין הויז וואס באגרעניצן ChatGPT נאר פאר פארשונג. ער האלט אז יעדער אפיס זאל אליין באשליסן.

אנאליזן און מיינונגען

- **1 יוני: קליינע מאדעלן און טורנינג** – דער BabyLM קאנקורס שטופט פאר קלענערע מאדעלן. AI21 Labs'ס טורנינג טעסט מיט מער ווי 10 מיליאן שמועסן און 1.5 מיליאן באטייליגטע: מען האט ריכטיג דערקענט א מענטש 73% פון די מאל, אבער א באט נאר 60%. ווייטער: וועגן דעם Shoggoth מעם.
- **2 יוני: CEOs** – Hamilton Nolan טענהט אז מען זאל אויטאמאטיזירן דעם CEO פאסטן.
- **5 יוני: AI קאוד** – טעק-פירער זארגן זיך אז צו פיל AI קאוד וועט ברענגען קאמפליצירטקייט און טעכנישע חובות. Bertrand Meyer טענהט אז AI העלפט נישט פאר פראגראמירער.
- **7 יוני: Andreessen** – Marc Andreessen האלט אז AI איז די גרעסטע אמצאה אלע מאל, און אז זי וועט "ראטעווען די וועלט".
- **8 יוני: מארק און פערזענליכע ערפארונג** – 14 טשארטס ווייזן ווי געלט און טאלאנט גיסן זיך אריין אין AI. דער ניוזלעטער פרעגט צי עכטער ביזנעס וועט נאכקומען. א איינצל-אונטערנעמער דערציילט אז ער האט פארגרעסערט זיין פארדינסט 8 מאל מיט ChatGPT.
- **9 יוני: "לערן AI שפעטער"** – א שרייבער האלט אז מען איז נאך אינמיטן דער רעשיגער תקופה, ענליך ווי ביים וועב און מאביל.

- **12 יוני:** **Stable Diffusion** און **LLMs** – Bloomberg געפינט אז **Stable Diffusion** פארגרעסערט ראסן-סטערעאטיפן מער ווי אין דער רעאליטעט. Icamtuf שרייבט וועגן די מענטשנגלייכע זייטן פון LLMs.
- **13 יוני:** **Noah Smith** – Noahpinion האלט אז מען דארף נישט "פורעמען" AI צו שיצן ארבעטער, נאר בויען אינסטיטוציעס וואס זאלן לאזן יעדן בליען.
- **16 יוני:** **מיינונגען** – MIT Technology Review שרייבט וועגן די געלעגנהייטן פאר AI אין אזיע. לויט ProductBoard ניצן 90% פון סטארטאפס AI. LeCun האלט אז AI איז נאך נישט קלוג ווי א הונט. לויט GitHub אנקעטע ניצן 92% פראגראמירער AI.
- **20 יוני:** **ארבעט און סכנות** – ווי AI טוישט ארבעט אין דריי אינדוסטריעס: פראפעסיאנעלע סערוויסעס, פילם און קאוד. ווי אזוי "עקזיסטענציעלע סכנה" איז געווארן דער גרעסטער מעם אין AI.
- **21 יוני:** **ארכיטעקטור** – a16z באשרייבט נייע ארכיטעקטורן פאר LLM אפליקאציעס. א שרייבער האלט אז מען זאל ניצן AI אלס א געצייג און נישט אלס אויטאמאטיזאציע.
- **22 יוני:** **די מענטשן הינטער AI** – דאטא-ארבעטער טוען איבערחזר'דיגע ארבעט פאר ווייניג געלט. לויט א שטודיע ניצן א היפשער טייל פון זיי אליין AI. ארבעטער אין עטיאפיע, עריטרעא און קעניע ווערן אויסגענוצט, און זיי פאראייניגן זיך יעצט אין יוניאנס.
- **23 יוני:** **Altman און דער מארק** – Time האט אינטערווירט Sam Altman. פאסטנס פאר ארבעט אין גענעראטיווע AI זענען געשטיגן מיט 20% אין מאי. פיטשעס געשריבן פון GPT-4 האבן 3 מאל מער שאנס צו באקומען געלט.
- **26 יוני:** **גרייס און ארבעט** – א נישט-באשטעטיגטער בריוו זאגט אז GPT-4 איז א מישונג פון קלענערע מאדעלן. דער וועג "גרעסער איז בעסער" גייט צו אן עק, און מען ווענדט זיך צו עפיציענץ. א שטודיע פון אייראפעאישע לענדער געפינט אז דער בום פון די 2010ער האט געשאפן פאסטנס. ווייטער: וועגן דער "לאסט פון וויסן".
- **27 יוני:** **פארשידנס** – אן אנקעטע פון 2016 וועגן פראגנאזן פון עקספערטן. א שטודיע פון 2,000 מענטשן וועגן האפענונג און פחד פון AI. גענעראטיווע AI פארפלייצט דעם וועב מיט אינהאלט פון נידעריגער קוואליטעט. בלאקטשעין מיט AI. Microsoft's Jared Spataro האלט אז AI פאדערט א נייעם צוגאנג צו ארבעט און פירונג.
- **28 יוני:** **קאוד-עקאנאמיע** – GitHub שאצט אז AI געצייג קענען שפארן ביז 30% פון א פראגראמירער'ס צייט, וואס קען ברענגען אן עקאנאמישע השפעה פון 1.3 טריליאן דאלאר אין אמעריקע.
- **29 יוני:** **מיינונגען** – Rohit Krishnan קלערט וואס עס וואלט געפאדערט צו בויען אן אגענט וואס פארבעסערט זיך אליין און קען דורכפירן וויסנשאפטליכע עקספערימענטן. Varun Shenoy האלט אז נאטירליכע שפראך איז נישט קיין גוטער אינטערפעיס. א שטודיע געפינט אז מענטשן גלייבן מער טוועטס פון GPT-3 ווי פון מענטשן.
- **30 יוני:** **צוקונפט** – a16z שרייבט אז דער קאמפיוטער וועט זיך טוישן פון א געצייג צו א באגלייטער. NFX האלט אז סטארטאפס מוזן ארבעטן שנעלער אין דער צייט פון AI.

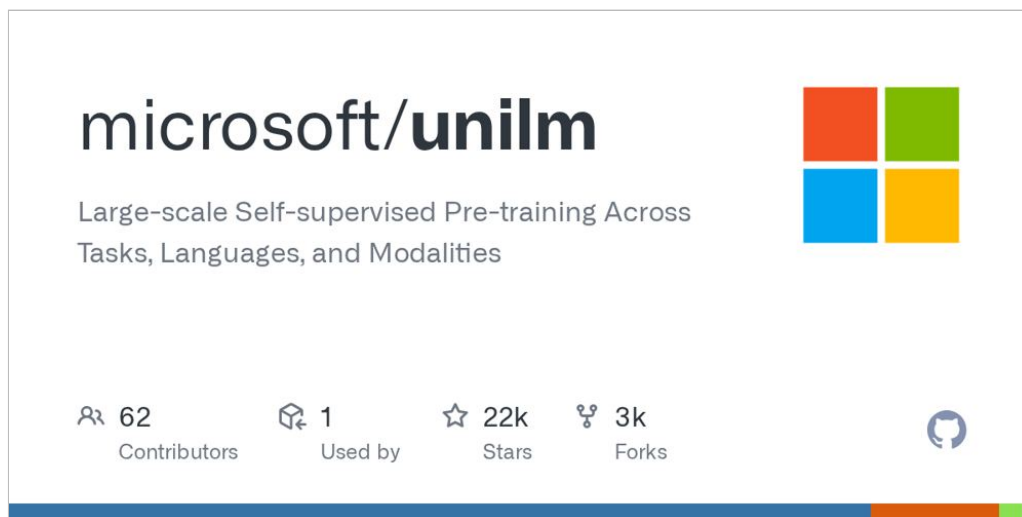
ווייטערע ידיעות

- **2 יוני:** טרו-קריים דיפפעינס – TikTok קאנטעס פארשפרייטן AI-קליפס פון קרבנות פון מארד, מערסטנס קינדער, וואס "דערציילן" ווי זיי זענען אומגעקומען.
- **5 יוני:** קורס און Minecraft – Google האט ארויסגעלאזט א קורס פון 9 טיילן אויף גענעראטיווע AI. Voyager, א באט אויף GPT-4, האט מיט הצלחה געשפילט Minecraft.
- **9 יוני:** דערקענען ChatGPT – אקאדעמיקער טענהן אז זייער אלגאריטם דערקענט ChatGPT פארשונג-ארבעטן מיט מער ווי 99% פינקטלעכקייט.
- **12 יוני:** לערנען וועגן AI – טעק-פירמעס, אוניווערסיטעטן און נאן-פראפיטס מאכן לעקציעס וועגן AI פאר סטודענטן, אבער ס'איז נאך נישטא קיין נאציאנאלע מסכמה וואס צו לערנען.
- **14 יוני:** Beatles און Roop – Paul McCartney האט געניצט AI צו אפשיידן John Lennon'ס קול פון א דעמא, צו מאכן א נייעם Beatles ליד. Roop שאפט דיפפעינס מיט איין קליק.
- **15 יוני:** דאקטוירים – דאקטוירים ניצן ChatGPT צו רעדן מער מיטפילנדיג מיט פאציענטן.
- **22 יוני:** Pixar און Marvel – Pixar האט געניצט Neural Style Transfer צו אנימירן דעם פייער-כאראקטער Ember אין "Elemental", און מיט GPUs האט זי דערגרייכט 20 מאל מער שנעלקייט. די עפענונגס-קרעדיטס פון "Secret Invasion" זענען געמאכט מיט AI.
- **28 יוני:** AI מעדיצין – א מעדיצין געשאפן מיט AI פון Insilico Medicine, קעגן א כראנישער לונגען-קרענק, וועט געטעסט ווערן אויף 60 מענטשן, נאכדעם וואס זי האט שוין מיט הצלחה דורכגעמאכט phase 1.

יולי 2023 איז געווען א חודש פון גרויסע מאדעל-לאנסירונגען: Meta האט ארויסגעלאזט Llama 2 מיט א קאמערציעלע ליצענץ און אין צוזאמענארבעט מיט Microsoft, Anthropic און Qualcomm האט לאנסירט Claude 2 מיט 100K טאקענס, און Stability AI האט ארויסגעלאזט SDXL 1.0. OpenAI האט געמאכט GPT-4 אלגעמיין צוגענגליך, געגרינדעט א "Superalignment" טיעם מיט 20% פון איר קאמפיוט, און איז געשטאנען אונטער דרוק: א FTC אויספארשונג, א קלאגע איבער קאפירעכט, און א פאל אין טראפיק. אין וואשינגטאן האבן גרויסע פירמעס זיך פארפליכטעט צו פרייוויליגע זיכערהייט-שריט, און די פירנדיגע לאבס האבן געגרינדעט א "Frontier Model Forum". אין דער פארשונג האט מען שטארק געארבעט אויף לאנגע קאנטעקסט (LongNet, Focused Transformer), "Lost in the Middle", אויף FlashAttention-2 און אויף נייע ארכיטעקטורן ווי RetNet.

מאדעלן און נייע לאנסירונגען

- **5 יולי: ChatGPT בלעטערן אפגעשטעלט** – OpenAI האט צייטווייליג אפגעשטעלט די "Browse with Bing" פונקציע אין ChatGPT, טראץ טענות פון ChatGPT Plus באניצער. די סיבה: די AI איז אומגעוואלט אריבערגעגאנגען צאל-מויערן (paywalls) און פריוואטקייט-איינשטעלונגען. OpenAI זאגט אז די פונקציע וועט צוריקומען נאך די נויטיגע פארריכטונגען.
- **6 יולי: Kosmos-2 ארויסגעגעבן** – דער UniLM טיעם ביי Microsoft, וואס האט אנטוויקלט איינעם פון די בעסטע מולטי-מאדאלע מאדעלן, האט ארויסגעגעבן דעמא-קאוד און וועיטס פאר Kosmos-2.



Kosmos-2 ארויסגעגעבן · 6 יולי 2023

- **7 יולי: GPT-4 פאר אלעמען** – OpenAI האט געמאכט GPT-4 אלגעמיין צוגענגליך, אזוי אז מען קען אריבערגיין פון GPT-3.5. אין דער זעלבער צייט האט InternLM ארויסגעלאזט אן אפן-סאורס טשעט-מאדאל פון 7 ביליאן פאראמעטערס, צוגעפאסט פאר פראקטישע באנוץ.
- **10 יולי: Google'ס מעדיצינישער טשעטבאט** – Google פרובירט אויס איר Med-PaLM 2 אין Mayo Clinic. דאס איז א ווערסיע פון PaLM 2 וואס ענטפערט אויף מעדיצינישע פראגעס. דער מאדעל האט

- געוויזן גוטע רעזולטאטן אין טראכטן, אין פארשטיין, און אין ענטפערס וואס דאקטוירים זענען מסכים מיט, הגם ער האט נאך עטליכע פעלערן אין גענויקייט. Google זאגט אז דער קונה'ס דאטא בלייבט פארשליסלט און Google האט נישט קיין צוטריט צו איר. Google לויפט דערמיט א פארמעסט קעגן Microsoft וועלכער וועט צוערשט ברענגען אזעלכע מאדעלן צו דאקטוירים.
- **10 יולי: Alibaba בילדער-מאכער** – Alibaba האט לאנסירט א געצייג וואס מאכט בילדער פון א טעקסט-באפעל.
 - **11 יולי: Salesforce קאוד-מאדעלן** – Salesforce האט ארויסגעלאזט קליינע קאוד-מאדעלן פון 7B פאראמעטערס. זיי שלאגן אלע מאדעלן פון דער זעלבער גרייס, און אפילו מאדעלן צוויי מאל אזוי גרויס.
 - **12 יולי: Claude 2** – Anthropic האט ארויסגעלאזט Claude 2, צוגענגליך דורך API און א בעטא וועבזייטל. דער מאדעל איז בעסער אין טראכטן, אין קאוד און אין חשבון, און גיט לענגערע ענטפערס. ער קען נעמען ביז 100K טאקענס אין יעדן באפעל, און דערמיט קען ער אנפאקן קאמפליצירטע טעכנישע אויפגאבעס און שרייבן לאנגע דאקומענטן.
 - **12 יולי: Baichuan-13B** – דער כינעזישער אנטרעפּרענער Wang Xiaochuan און זיין סטארטאפ Baichuan Intelligence האבן לאנסירט Baichuan-13B, א מאדעל וואס זאל קאנקורירן מיט OpenAI.
 - **13 יולי: Lince Zero אויף שפאניש** – דער שפאנישער סטארטאפ Clibrain האט לאנסירט Lince Zero, אן אפן-סאורס מאדעל וואס איז אויסגעארבעט פאר דער שפאנישער שפראך.
 - **14 יולי: Orca קאפיס** – מען האט אפן ארויסגעגעבן א 13B ווערסיע וואס איבערמאכט די Orca ארבעט. Orca איז א "דיסטילאציע" פון GPT-4: מען לערנט א קלענערן מאדעל מיט קלוגע באפעל-טעכניקן. דערנאך איז אויך געקומען OpenChat Orca, וואס ניצט אויך GPT-4'ס ערקלערונגען פונקט ווי Orca.
 - **14 יולי: Stable Doodle** – Stability AI האט לאנסירט א סערוויס וואס פארוואנדלט א שנעלע צייכענונג אין א פולן בילד.
 - **17 יולי: Meta טעקסט-בילד מאדעל** – Meta האט ארויסגעלאזט א decoder-only מאדעל וואס באהאנדלט טעקסט און בילדער אויסגעמישט ווי טאקענס. דערמיט קען ער מאכן בילדער פון טעקסט און טעקסט פון בילדער, און ער קען אפילו ריכטיג צייכענען הענט.
 - **18 יולי: NotebookLM** – Google האט ארויסגעלאזט NotebookLM, אן עקספערימענטאלע AI אפליקאציע פאר נאטיצן.
 - **19 יולי: Llama 2** – Meta האט ארויסגעלאזט Llama 2 מיט א קאמערציעלע ליצענץ. דער מאדעל איז טריינירט אויף 2 טריליאן טאקענס, ביז 70B פאראמעטערס. די פארשונג-ארבעט באשטעטיגט אז ריכטיג טריינירן דעם "reward model" איז די יסוד-זאך, און לויטן קאפ פון RLHF ביי Hugging Face איז RLHF געווען דער שליסל, אויסער דער פארדאפלטער טריינינג-דאטא. Microsoft שטיצט Llama 2 אויף Azure און Windows, און Qualcomm ברענגט אים צו סמארטפאונס. לויט Matt Rickard האט Meta געעפנט דעם מאדעל כדי צו שוואכן קאנקורענטן מיט פריוואטע מאדעלן, און אלס א פרייע מדרגה צו Meta'ס אייגענע קאמערציעלע אנגעבאט.
 - **19 יולי: MPT מיט 8K** – Mosaic האט ארויסגעלאזט MPT מאדעלן מיט 8K קאנטעקסט. זיי זענען צוגעטריינירט געווארן אויף נאך 500 ביליאן טאקענס, 3 טעג אויף H100s, און זענען גוטע באזע-מאדעלן פאר fine-tuning.

- **19 יולי:** **GPT-4 בילדער צוריקגעהאלטן** – OpenAI האלט צוריק דאס בילדער-דערקענען פון GPT-4 פון ברייטן ציבור, ווייל מען האט מורא אז ער וועט קענען אידענטיפיצירן ספעציפישע מענטשן.
- **20 יולי:** **Apple GPT** – Apple בויט א פריימווערק Ajax און א טשעטבאט "Apple GPT" כדי צו קאנקורירן אין דעם געביט. דאס איז נאך עקספערימענטאל, אבער אינסיידערס ערווארטן א גרויסע מעלדונג נעקסטן יאר.
- **21 יולי:** **ChatGPT און Copilot** – OpenAI האט צוגעלייגט "custom instructions" אין ChatGPT, אזוי אז מען דארף נישט יעדעס מאל איבערשרייבן די זעלבע אנווייזונגען. Copilot Chat'ס GitHub איז ארויס אין א באגרעניצטער פובליק בעטא פאר פירמעס.
- **24 יולי:** **FreeWillie** – Stability AI האט ארויסגעגעבן א סעריע מאדעלן וואס זענען צוגעפאסט צו פאלגן אנווייזונגען, נאר פאר פארשונג.
- **25 יולי:** **טשעטבאטס ווערן ברייטער** – ChatGPT קומט צו Android און מען קען זיך שוין פאררעגיסטרירן. Bard איז צוגענגליך אין נייע שפראכן און לענדער, און מען קען אריינלייגן בילדער. אבער א פארשונג אין 13 פארשידענע מצבים האט געוויזן אז Bard איז נאך שוואך אין פארשטיין בילדער. Bing Chat ווערט אויסגעפרובירט אויף Chrome און Safari, אפשר מיט באגרענעצונגען, און מען טעסט אויך א דונקעלע טעמע און אן איינבוי פאר פירמעס.
- **25 יולי:** **לענגערער Llama 2 קאנטעקסט** – Nous Research האט אויסגעברייטערט Llama-2-13B ביז 8K טאקענס דורך positional interpolation. Together, וואס האט לעצטנס אויפגענומען \$20 מיליאן, האט אויסגעברייטערט Llama 2 7B ביז 32K, אבער זיי ווארענען אז דאס "פארגעסן די מיטן" איז נאך אלץ א פראבלעם.
- **25 יולי:** **Cerebras 3B** – Cerebras, וואס בויט טשיפס צו טריינירן מאדעלן, האט טריינירט א 3B מאדעל אויף דעם SlimPajama דאטאסעט. דער מאדעל איז גאנץ שטארק און קען לויפן אויף מאביילפאונס.
- **28 יולי:** **SDXL 1.0** – Stability AI האט לאנסירט Stable Diffusion XL 1.0, צוגענגליך אויף Amazon Bedrock און דורך איר אייגענעם API. ער האט בעסערע קאלירן און דעטאלן, און א בעטא אפציע צו fine-tune פאר ספעציעלע בילדער.
- **28 יולי:** **מאדעל פאר טאבעלעס** – א נייער מאדעל טריינירט אויף 500B טאקענס טאבעלע-דאטא. ער קען אויספילן, מאכן און באלאנסירן סטרוקטורירטע דאטא, אפילו ווען מען האט גארנישט קיין דאטא (zero-shot).

פארשונג און טעכניק

- **3 יולי:** **Hyena פאר DNA** – דער state space מאדעל Hyena, מיט 1 מיליאן טאקענס קאנטעקסט, איז בעסער געווארן אין DNA 23 אויפגאבעס.
- **3 יולי:** **בעסערע CLIP** – CLIPA-v2 ניצט קירצערע סעקווענצן און דערגרייכט 81.1% zero-shot ImageNet פאר בלויז \$10,000. אן אנדער פראיעקט פילטערט וועב-דאטא פאר CLIP און דערגרייכט די בעסטע zero-shot רעזולטאטן.
- **3 יולי:** **בילד- און 3D פארשונג** – עטליכע ארבעטן: MVDiffusion מאכט פון איין בילד א פאנאראמע אדער פארשידענע אנזיכטן. FGD פארבעסערט בילד-צו-בילד איבערזעצונג אן נאכן טריינירן. SITA-

- VAE און ASLDM מאכן 3D פארמען פון בילדער אדער טעקסט. - א מעטאד מאכט גלאטע ווידעאס פון טעקסט מיט שליסל-פריימס. - WavePaint פילט אויס פעלנדע טיילן פון א בילד שנעלער. - latent editing מיט DDPM קען פארוואנדלען א פנים, למשל אין Tom Cruise.
- **3 יולי: מעדיצינישע 3D בילדער** - "Volume Fusion" טריינירט 3D מעדיצינישע מאדעלן אן גרויסע מאסן באצייכנטע דאטא, און פארבעסערט די גענויקייט פאר פארשידענע טיילן פונעם גוף.
 - **4 יולי: Classifier-Free guidance** - לויט א פארשונג-ארבעט קען א נייע guidance-מעטאד ביים דעקאדירן פארדאפלען דעם עפעקטיוון גרייס פון א מאדעל, אן נאכאמאל טריינירן.
 - **4 יולי: דערקענען דיפפעינס** - DADF ניצט SAM מיט אן adapter און RGA צו דערקענען פארפעלשטע פנימער און געפינען גענוי וואו זיי זענען פארפעלשט.
 - **4 יולי: בילדער פון מוח** - פארשער ניצן Stable Diffusion צו רעקאנסטרווירן בילדער פון fMRI מוח-אקטיוויטעט, אן נאכן טריינירן. MindEye פון Stability AI איז דער נייסטער שפיץ אין דעם געביט.
 - **4 יולי: 3D באוועגונג** - א מעטאד גיט מער רעאליסטישע פאזעס פון 3D באוועגונג אין פארבינדונג מיט דער סביבה.
 - **4 יולי: MEGABYTE** - א פאדקאסט מיטן מחבר פון Meta's MEGABYTE. די ארבעט באזייטיגט טאקעניזאציע, אזוי אז איין מאדעל קען ארבעטן מיט יעדע סארט דאטא.
 - **5 יולי: ריזיקע פון זעלבסטמארד** - א מאדעל פרובירט פאראויסזאגן די ריזיקע ביי בייפאלאר פאציענטן, לויט די סימפטאמען וואס שפיגלען זיך אפ אין זייערע פאסטן.
 - **5 יולי: פאלינאמען מיט RL** - מען פארוואנדלט די רעזאלוציע פון סינגולאריטעטן אין א שפיל און לייגט צו reinforcement learning, מיט שטארקע רעזולטאטן.
 - **5 יולי: Hypervectors** - א נייער צוגאנג ניצט גרויסע וועקטארן וואס שטעלן פאר באגריפן, כדי צו טראכטן מער אבסטראקט און עפעקטיוו.
 - **6 יולי: Zero-shot בילד-טעקסט** - א טעכניק וואס קאנצענטרירט זיך אויף פיינע דעטאלן און שליסל-ווערטער.
 - **6 יולי: לאנגע קאנטעקסט** - LongNet ניצט dilated attention אויף 1 ביליאן טאקענס, אן צו פארלירן ביי קורצע סעקווענצן. Focused Transformer און LongLlama ניצן contrastive learning און וואקסן ביז עטליכע הונדערט טויזנט טאקענס. "Lost in the Middle" ווייזט אז מאדעלן ארבעטן שוואך מיט אינפארמאציע אין דער מיטן פונעם קאנטעקסט. L-Eval גיט איבער 2,000 פראגע-ענטפער פארן אפשאצן לאנגע קאנטעקסט.
 - **6 יולי: SAM-DA און WebGLM** - SAM-DA פארבעסערט נאכט-טראקינג פון דראונס מיט ווייניגער רעסורסן. WebGLM ענטפערט פראגעס דורכן זוכן אויפן וועב, ביליג און גענוי.
 - **7 יולי: רעדאגירן בילדער** - DragonDiffusion רוקט און טוישט אביעקטן אין א בילד. FreeDrag פארבעסערט Subject-Diffusion. DragGAN פערזענליכט בילדער פון איין בילד. BoxDiff לאזט באשטימען וואו אביעקטן שטייען. TF-ICON שטעלט אריין אביעקטן אין בילדער. אלע ארבעטן אן נאכאמאל טריינירן.

- **7 יולי:** סינטטישע דאטא – א דעבאטע צי מאדעלן וואס מאכן אליין טריינינג-דאטא איז א פארבעסערונג אדער אן אילווע. איין מיינונג איז אז עס ארבעט ווי א פילטער. פירמעס פרובירן שוין צו מאכן דאטא פון אנפאנג.
- **7 יולי:** HIPIE און FLACUNA – HIPIE איז א מעטאד פאר סעגמענטאציע פון בילדער אויף פארשידענע מדרגות. FLACUNA איז VICUNA וואס איז fine-tuned מיט FLANMINI, כדי צו לייען פראבלעמען בעסער.
- **10 יולי:** MoE און RL – Mixture of Experts מאדעלן לאזן זיך אפשר לייכטער צופאסן מיט instruction tuning. Inverse RL אן דער אינעווייניגסטער RL שליף איז שנעל און שטארק. מאדעלן קענען בלויז באגרעניצט טראכטן "וואס וואלט געווען אויב".
- **10 יולי:** TTA בענטשמארק – א בענטשמארק פאר Test-Time Adaptation מעטאדן.
- **11 יולי:** INT-FP-QSim און DPC-DUN – INT-FP-QSim איז אן אפן-סאורס סימולאטאר וואס פרובירט מאדעלן ביי נידעריגערע פרעציזיע. DPC-DUN פארבעסערט בילד-רעקאנסטרוקציע אין Compressive Sensing.
- **12 יולי:** CoactSeg און TIM – CoactSeg פארבעסערט דאס געפינען נייע MS שאדנס מיטן נייעם MS-23v1 דאטאסעט. TIM לערנט איבערזעצונג מיט ריכטיגע און פאלשע ביישפילן, און איז בעסער אויף WMT2022.
- **13 יולי:** CAT-ViL, אנטאלאגיע, סינעמאגראף – CAT-ViL ענטפערט פראגעס אויף אפעראציע-ווידעאס פאר סטודענטן. סטרוקטורירטע באפעלן לויט אן אנטאלאגיע פארבעסערן ChatGPT, ווי מען האט געוויזן אויף אוקראיניש. Text2Cinemagraph מאכט באוועגליכע בילדער פון טעקסט.
- **14 יולי:** gzip און קול – gzip איז די בעסטע מעטאד צו דערקענען זאצן וואס פאסן נישט צו דער אויפגאבע, אן טייערן טריינירן. א פשוטע מעטאד פארוואנדלט א קול דורך nearest neighbors.
- **14 יולי:** וועמענס מיינונגען? – Anthropic פארשט מיט וועלכע מענטשן אין דער וועלט די ענטפערס פון מאדעלן שטימען מערסטנס, ווייל די מאדעלן שפיגלען נישט אפ גלייך די מיינונגען פון דער מענטשהייט.
- **17 יולי:** אפטימיזאציע און אנימאציע – מען האט באוויזן אז א מאל א גרויסער שפרונג ביים learning rate גיט שנעלערע קאנווערגענץ. AnimateDiff לייגט צו באוועגונג צו בילדער פון פערזענליכע מאדעלן. Drunkard's Dataset העלפט באשטימען קאמערע-באוועגונג אין א טוישנדיגער סביבה.
- **18 יולי:** מאראל – א פארשונג-ארבעט ווייזט אז מאדעלן זענען שוואך אין מאראלישע אויפגאבעס, קעגן דער טענה אז מאראל וואקסט מיט די פעאיגקייטן.
- **18 יולי:** FlashAttention-2 – די צווייטע ווערסיע פון FlashAttention, מיט א געזאגטער 5-9 מאל בעסערער ליסטונג קעגן געווענליכן attention.
- **18 יולי:** נאך ארבעטן – א retriever וואס לערנט אויסקלייבן די בעסטע ביישפילן פאר LLMs. mBLIP, דער ערשטער פיל-שפראכיגער בילד-שפראך מאדעל. א סיסטעם וואס פארוואנדלט מעשיות אין ווידעאס. Video-FocalNet פאר דערקענען ווידעאס.

- **19 יולי: RetNet** – Microsoft's UniLM האט ארויסגעלאזט RetNet, א מיטלוועג צווישן רעקורענטע מאדעלן און טראנספארמערס. מאדעלן פון 7B, טריינירט אויף 100B טאקענס, זענען בעסער ווי טראנספארמערס. מען האט טריינירט אויף 512 AMD MI200 GPUs.
- **19 יולי: נאך ארבעטן** – Skeleton Sequence Learning פאר דערקענען אקציעס. COLLIE פאר טעקסט מיט קאמפליצירטע כללים. הארמאניזאציע פון בילדער אן אסאך דאטא. Neural Video Depth Stabilizer מיטן דאטאסעט Video Depth in the Wild.
- **20 יולי: GPT-4 טוישט זיך?** – א פארשונוג געפינט אז GPT-3.5 און GPT-4 טוישן זייערע רעזולטאטן נאך דערהיינטיגונגען. אנדערע ווידערהייגן די טענה אז GPT-4 ווערט ערגער, און זאגן אז דאס איז א טויש אין פירונג נאך fine-tuning, נישט קיין פארלוסט אין פעאיגקייט.
- **20 יולי: וויזואלע ארבעטן** – SHARCS לערנט פון פארשידענע סארטן דאטא אויף איינמאל. SAM-PT טראקט אביעקטן אין ווידעאס. RepViT איז א לייכטע נעץ. א גרויסער דאטאסעט פון מיליאנען פריימס מיט הענט וואס רירן אן אביעקטן. OnlineRefer געפינט אביעקטן אין ווידעאס לויט אן אנווייזונג.
- **21 יולי: ICML** – Google DeepMind שטעלט פאר איבער 80 פארשונוג-ארבעטן ביי ICML, אריינגערעכנט AlphaFold, פיוזשאן, PaLM-E און Phenaki. ICML האט אויך ארויסגעגעבן איר רשימה פון די בעסטע ארבעטן, אויך פון Apple.
- **21 יולי: אפטימיזאציע און חשבון** – א לערנבוך איבער Amortized Optimization. א פרי אלגאריטם וואס מאכט גרויסע מאטריצן-מולטיפליקאציע נאר מיט אדיציע, וואס קען שפארן אויף טשיפס.
- **24 יולי: אנגענטן און וויסן** – א שפראך-מאדעל אלס פלאנירער פארבעסערט אן אנגענט. RAG העלפט מאדעלן וואס גלייבן צופיל אין זיך צו דערקענען זייערע גרענעצן. SlowTV פון YouTube פארבעסערט דאס שאצן פון טיפקייט.
- **25 יולי: זכרון און attention** – RecallM פאר לאנג-טערמיניגן זכרון. א בלאג טענהט אז attention איז "off by one" און פארשלאגט א פארריכטונג. AlignDet פאר דערקענען אביעקטן.
- **26 יולי: וויזואלע ארבעטן** – מעדיצינישע בילדער פארבונדן מיט באריכטן. CTVIS פאר ווידעא-סעגמענטאציע. Described Object Detection לויט פרייע באשרייבונגען.
- **27 יולי: פריוואטע באפעלן** – מיט multiparty communication האט מען לויפן געלאזט Llama 7B אן אנטפלעקן דעם באפעל. דאס האט געדויערט בערך 5 מינוט פאר 1 טאקען, אבער דאס איז די ערשטע מאל וואס עס האט טאקע געארבעט.
- **27 יולי: נאך ארבעטן** – Graph Neural Networks צו כאפן געלט-וואשן. SwinMM פאר מעדיצינישע בילדער. FacTool כאפט פאקטישע פעלערן. SgMg פאר ווידעא-סעגמענטאציע. מאדעלן ווייזן אז מוח-כוואליעס העלפן אפשר זכרון און חשבון גלייכצייטיג.
- **28 יולי: נאך ארבעטן** – CFSum פאר מולטי-מאדאלע סומארן. GoodsAD פאר אנאמאליעס אין סופערמארקעטן. HQTrack, צווייטער פלאץ אין JourneyDB, VOTS2023, איבער 4 מיליאן בילד-טעקסט פארן פון Midjourney.
- **31 יולי: סקילס, ראבאטן** – א געהעריג אויסגעקליבענער דאטאסעט גיט מיט 1 ביליאן טאקענס די זעלבע רעזולטאטן ווי 3 ביליאן. Google's RT-2 ניצט טראנספארמערס פאר ראבאטן. Thinker לערנט פלאנירן אין א סימולאציע, Sokoban און Atari. vox2vec פאר CT בילדער.

געצייג פאר פראגראמירער און אפן-סאורס

- **3 יולי:** **GPT-Migrate** און **רעטענישן** – GPT-Migrate איבערפירט קאוד פון איין שפראך אדער פריימווערק צו אן אנדערע. Sasha Rush האט רעטענישן וואס לערנען מאדעל-פאראלעליזם און gradient accumulation.
- **5 יולי:** **געצייג** – S.A.T.U.R.D.A.Y פאר קול-אינטערפייסעס. Aider, קאוד מיט GPT פון דער קאמאנד-ליין. Scenic פאר ווערק פון attention אין בילדער. א לערנבוך איבער learning theory.
- **6 יולי:** **געצייג** – א vector database מיט Metal shaders, פיל שנעלער פאר ChatGPT פלאגיןס. Streamlit מיט LLMs. א ווידעא איבער לינעארע NanoT5 attention. טריינירט 150 מאל שנעלער.
- **7 יולי:** **LlamaIndex 0.7.0** – מער מאדולאר, מיט אבסטראקציעס פאר OpenAI, HuggingFace און PaLM, און בעסערע מעטאדאטא.
- **10 יולי:** **געצייג** – ChatGPT.js. InstructBLIP פון Salesforce אויף Hugging Face. א פאדקאסט פון Replit AI.
- **11 יולי:** **געצייג** – Haystack, אן אפן-סאורס פריימווערק מיט 9,335 שטערן. Danswer ענטפערט פון אינערליכע דאקומענטן מיט ציטאטן. Alibaba'ס Chat2DB, א SQL קליענט מיט ChatGPT. GGML מיט MPI פרוברט לויפן לאזן Llama 65B אויף Raspberry Pis.
- **12 יולי:** **4-בייט** – bitsandbytes גיט 4-בייט inference מיט ביז 4 מאל שנעלערקייט, פשוט צו ניצן מיט HuggingFace. Jumanji, RL סביבות אין gpt-prompt-engineer. JAX. וואס פרוברט און שאצט באפעלן.
- **13 יולי:** **אגענטן** – GPT Researcher, אן אוטאנאמער פארשונג-אגענט. Reflexion, קאוד פאר אגענטן מיט מינדליכן RL.
- **14 יולי:** **PR-Agent** – אן אפן-סאורס געצייג פון CodiumAI וואס איבערקוקט pull requests, שרייבט באשרייבונגען און פארשלאגט קאוד. Aeon פאר time series.
- **17 יולי:** **געצייג** – א קריטיק אויף Microsoft'ס LangChain. FastEdit'ס Semantic Kernel. צו אריינלייגן נייע וויסן מיט איין באפעל. אן אומאפיציעלער Claude API. א פארגלייך פון perplexity ביי quantized מאדעלן. Petals מיט PEFT.
- **18 יולי:** **MetaGPT** – MetaGPT מאכט פון איין שורה דאקומענטן, APIs און אנאליזן. DemoGPT מאכט אן אפליקאציע פון א באפעל. AutoChain, א לייכטער אגענט-פריימווערק.
- **20 יולי:** **Keras Core** – א פארקוק אויף Keras 3.0, וואס לויפט אויף JAX, TensorFlow און PyTorch.
- **21 יולי:** **געצייג** – Ollama לויפט מאדעלן אויף macOS, א פרייוואטער ChatGPT אין דיין VPC'ס AXLearn אויף Jax מיט פאראלעליזם דורכן קאמפיילער. OpenBB מיט LlamaIndex. Mojo קעגן C++.
- **24 יולי:** **LangSmith** – LangChain'ס פלאטפארמע צו דעבאגן, טעסטן און אכטונג געבן אויף אפליקאציעס. Copilot פארריכט Jupyter. Andrej Karpathy האט געשריבן Llama 2 אין ריין C, מיט הילף פון GPT-4.
- **27 יולי:** **קורס** – א פרייער קורס פון Hugging Face פאר טראנספארמערס מיט קול.

- **28 יולי:** GrammarGPT – פאר כינעזישע גראמאטיק. ReliableGPT פאר העכערע זיכערקייט ביים OpenAI API.
- **31 יולי:** TGI – Hugging Face'ס Text Generation Inference סערווער, טוישט די ליצענץ אזוי אז מען טאר אים שוין נישט ניצן קאמערציעל. Khoj, א פערזענליכער אסיסטענט.

פירמעס, געלט און ביזנעס

- **3 יולי:** Casetext און MosaicML – \$650 מיליאן פאר Casetext, וואס האט געדויערט א צענדליג יאר צו בויען, און \$1.3 ביליאן פאר MosaicML.
- **4 יולי:** Runway און Playground – RunwayML האט אויפגענומען \$141 מיליאן, מערסטנס פאר פארשונג. Playground AI האט אויפגענומען \$40 מיליאן.
- **5 יולי:** SoftBank – יאגט זיך נאך, טראץ \$140 ביליאן שוין אינוועסטירט.
- **7 יולי:** ChatGPT פאלט – דער טראפיק איז געפאלן 9.7% פון מאי ביז יוני (10.3% אין אמעריקע). Adobe באגרענעצט זיינע ארבעטער פון ניצן דרויסנדיגע AI געצייג, פונקט ווי Amazon, Apple און Samsung.
- **10 יולי:** Mistral – דער מעמא וואס האט אויפגענומען €105 מיליאן פאר א סטארטאפ פון בלויז פיר וואכן.
- **13 יולי:** x.AI – Elon Musk האט לאנסירט x.AI, מיטן ציל "צו פארשטיין דעם אוניווערס". ער האט צוגעזאגט פיינשאפטן אין דער פירמע, וואס ער האט געטענהט איז שוין ווערט צענדליגער ביליאנען.
- **13 יולי:** דאטא-אפמאכן – Shutterstock ברייטערט אויס איר אפמאך מיט OpenAI: דאטא פאר טעכנאלאגיע. AP האט אויך געמאכט אן אפמאך מיט OpenAI.
- **14 יולי:** Shopify, Stability – Shopify'ס Sidekick, אן אסיסטענט פאר געשעפט-אייגנטימער. א מיטגרינדער פון Stability AI טענהט אז מען האט אים אפגענארט צו פארקויפן זיינע 15% פאר \$100.
- **18 יולי:** Wix – מאכט וועבזייטלעך פון א באפעל.
- **20 יולי:** Superhuman, McKinsey – Superhuman האט צוגעלייגט AI אין אימעילס. McKinsey ארבעט צוזאמען מיט Cohere.
- **26 יולי:** Neura און Cohere – Neura Robotics האט אויפגענומען \$55 מיליאן, און ערווארט \$450 מיליאן באשטעלונגען אין 5 יאר. Scale'ס LLM Engine פאסט צו Cohere'ס Llama 2. Coral פאר פירמעס. Spotify'ס CEO רעדט וועגן פערזענליכקייט און סומארן פון פאדקאסטן.
- **27 יולי:** OpenAI אדוואקאטן – גרויסע אדוואקאטן פון גרויסע פירמעס גייען איבער צו OpenAI.
- **31 יולי:** OverflowAI – Stack Overflow לאנסירט AI, אלפא אין אויגוסט, מיט VS Code און Slack.
- **3 יולי:** עטליכע נייע אפליקאציעס – Editby (YouTube טראנסקריפציע), Wondercraft (פאדקאסטן), BrewNote און Circleback (נאטיצן פון זיצונגען), Myreader (פראגעס אויף ביכער), MindOS (אגענטן), Coframe (A/B טעסטן), Parse, Swap Anything, Spoke, PhantomBuster, ChatCSV, Codejet (Figma צו קאוד), Eden AI, FixMeBot, Wope, Seodity, ElevateHQ, Poster Studio, iPhone, Tidepool, Lingoedit, AI Logo Art פאר Lindo, ProApp, Prompt Pup, Rewind

רעגירונגען, געזעצן און זיכערהייט

- **3 יולי: EU AI Act** – גרויסע אייראפעאישע פירמעס קריטיקירן דעם געזעץ אז ער איז נישט עפעקטיוו און שאדט קאנקורענץ. דער געזעץ, דער ערשטער אויף דער וועלט, דעפינירט "foundation models" און לייגט פליכטן אויף הויך-ריזיקע סיסטעמען. א פארשונג פון Stanford ווייזט אז רוב מאדעל-פירמעס פאלגן נאך נישט אינגאנצן נאך די אייראפעאישע פריוואטקייט-כללים.
- **4 יולי: OpenAI שמועסן** – OpenAI באריכטעט וואס עס האט געלערנט פון שמועסן איבער דער וועלט.
- **4 יולי: Google דאטא** – Google האט געטוישט איר פריוואטקייט-פאליסי, אזוי אז זי קען ניצן אלע פובליקע אינפארמאציע צו טריינירן AI, און דאס ברענגט זארגן איבער פריוואטקייט און קאפירעכט.
- **6 יולי: OpenAI – Superalignment** – OpenAI ניצט 20% פון איר קאמפיוט פאר דעם ענין פון אליינמענט פון מאדעלן וואס זענען קלוגער פון מענטשן, און דינגט אויף אסאך מענטשן.
- **10 יולי: אמעריקאנער מיליטער** – די פענטאגאן טעסט פינף LLMs, מיט Scale AI.
- **11 יולי: קלאגעס** – א דריי שרייבער האבן געקלאגט OpenAI און Meta אז זייערע מאדעלן זענען אומגעזעצליך טריינירט אויף קאפירעכט מאטעריאל. איבער 8,500 שרייבער פאדערן ערלויבעניש און באצאלונג.
- **14 יולי: FTC** – די FTC פארשט אויס OpenAI איבער פערזענליכע אינפארמאציע.
- **21 יולי: פארפליכטונגען** – Microsoft, Google און OpenAI פארפליכטן זיך פרייוויליג צו שריט פון די ווייסע הויז, ביז קאנגרעס וועט מאכן געזעצן.
- **24 יולי: OpenAI** – דער קאפ פון trust and safety גייט אוועק נאך א יאר און א האלב.
- **27 יולי: Frontier Model Forum** – פירנדע לאבס האבן געגרינדעט א פארום פאר זיכערהייט.
- **27 יולי: יאפאן** – METI וועט בויען א סופערקאמפיוטער, 2.5 מאל שטארקער, פאר היימישע פירמעס.

אנאליזן און מיינונגען

- **3 יולי: AI אינזשעניר** – Latent Space באשרייבט דעם אויפשטייג פון "AI אינזשעניר". א שרייבער האט אויספרובירט 200 געצייג און איז געווען אנטוישט. Matt Rickard פרעגט צי דער קרייז צווישן AI פירמעס און הארדווער-פירמעס איז געזונט.
- **4 יולי: ארבעט** – Code Interpreter קען איבערמאכן פינאנצן, לויט Every. Ben Evans שרייבט איבער אויטאמאציע פון ארבעט. א דיסקוסיע צי טעכנאלאגיע-מענטשן זענען אין אן "עכא-קאמער". ארבעטער אין טעכנאלאגיע לערנען זיך שנעל AI.
- **5 יולי: גרויסע וויסאגונגען** – Zhengdong Wang טענהט אז די וואקסונג פון די ווירטשאפט פער קאפ איז נאך קיינמאל נישט אריבערגעשטיגן Scott Alexander 3%. באשרייבט פארשידענע מעגליכקייטן. Freddie DeBoer ווארנט פון גאווה.
- **6 יולי: וועיטס** – Ryan Moulton שרייבט וועגן ווי דיגיטאלע מוחות קענען וויסן. א טענה אז וועיטס זענען נישט באמת "אפן סאורס".

- **7 יולי: AI זיכערהייט** – א טענה אז די ווערטער "שפאלטונג" זענען אהיסטאריש. א מענטש-צענטרירטער צוגאנג.
- **11 יולי: Hassabis** – אן אינטערוויו מיט Demis Hassabis. AGI פון ערשטע פרינציפן. Davidson איבער ווי שנעל AGI קען קומען. Techdirt ווארנט אז רעגולאציע שטארקט די גרויסע פירמעס. סימבאלישע רעגירעסיע פאר וויסנשאפט.
- **12 יולי: מיינונגען** – John Luttig, Jeremy Howard טענהט קעגן שטרענגע רעגולאציע. Nick Clegg טרייבט Meta'ס אפן-סאורס. AMD פרובירט קאנקורירן מיט Nvidia.
- **13 יולי: Gates** – Bill Gates זאגט אז די ריזיקעס זענען אמת אבער מען קען זיי באהערשן. א פראפיל פון Anthropic. די בויסטיינער פון גענעראטיווע AI. Mike Loukides: מען דארף מורא האבן נאר פון נישט אריבערגיין צו AI פראגראמירן.
- **14 יולי: געזונט** – א באריכט איבער 145 פירמעס מיט \$20 ביליאן פינאנצירונג און 47,000 ארבעטער.
- **17 יולי: דעטעקטארן** – AI דעטעקטארן האלטן אז די אמעריקאנער קאנסטיטוציע איז געשריבן געווארן דורך AI. Bain זוכט געלעגנהייטן פאר סטארטאפס.
- **18 יולי: a16z** – a16z טענהט אז וואשינגטאן זאל נישט שטערן AI אין געזונט. א פאטענט פאר AI. 2/3 פון שאפער ניצן שוין AI.
- **19 יולי: אפליקאציעס** – פראבלעמען ביים בויען AI אפליקאציעס. JPMorgan איבער פירונג. Stephen Wolfram. Fable Studios האט געוויזן א טעלעוויזיע-עפיזאד געמאכט מיט AI אינמיטן א סטרייק.
- **20 יולי: ריזיקעס** – Noema טענהט אז די ריזיקע פון אויסשטארבן איז איבערגעטריבן. צי מאשינען קענען פילן ווייטאג.
- **24 יולי: Andreessen** – Marc Andreessen ניצט ChatGPT צו לערנען זיין זון. 13 פירמעס צו באאכטן. "בלוי-קאלנער" AI ארבעט.
- **25 יולי: גרובן** – Packy שרייבט איבער שוץ-גרובן אויסער דעם פראדוקט אליין.
- **26 יולי: Oppenheimer מאמענט** – א שרייבער טענהט מען זאל נישט אפשטעלן AI וואפן, נאר זיי פירן מיט שטארקע כללים. אמעריקאנער שטילשטאנד.
- **27 יולי: a16z** – AI שפילן וואס ענדיגן זיך קיינמאל נישט, פערזענליכע פינאנצן, און Merchant of Record.
- **28 יולי: TinyML** – פארוואס TinyML איז שווער. Azeem Azhar וועגן א "צווייטער מוח" מיט Notion.
- **31 יולי: ארבעט** – McKinsey איבער ארבעט. דעטעקטארן ארבעטן נישט קעגן נארן אין שול. Tomasz Tunguz: דאטא-טיעמס וועלן מער באדארפן ווערן.

ווייטערע ידיעות

- **3 יולי: רעליגיע** – The Walrus שרייבט איבער מענטשן וואס באהאנדלען AGI ווי א רעליגיעזע כח.
- **6 יולי: באקאנטע פערזענליכקייטן** – באקאנטע מענטשן מאכן AI קאפיעס פון זיך פאר רעקלאמעס.
- **7 יולי: סקענס** – CNET שרייבט וועגן פולע גוף-סקענס מיט AI.

- **10 יולי: Wimbledon** – באטראכט צו פארביטן די ליניע-ריכטער מיט AI. Christopher Nolan זעט Oppenheimer'ס געשיכטע אלס א ווארענונג. קאזאלע AI.
- **11 יולי: דעטעקטארן** – AI דעטעקטארן באשולדיגן אומגערעכט מענטשן וואס ענגליש איז נישט זייער מוטערשפראך.
- **21 יולי: Genesis** – Google פרובירט Genesis, וואס שרייבט נייעס-ארטיקלען.

אויגוסט 2023 איז געווען א חודש פון שטארקע אפן-סאורס מאדעלן פאר קאוד: Meta האט ארויסגעלאזט Code Llama, און באלד דערויף האבן Phind און WizardLM ארויסגעגעבן פארבעסערטע ווערסיעס וואס זענען נאענט צו GPT-4 אדער האבן אים איבערגעיאגט אויף געוויסע בענטשמארקס. OpenAI האט ערלויבט צו טריינירן gpt-3.5-turbo אויף אייגענע דאטא, לאנסירט ChatGPT Enterprise, און ווערט געשאצט אז זי וועט באלד דערגרייכן א ביליאן דאלאר הכנסות א יאר. אויף דער געלט-זייט האט Hugging Face אויפגעשטעלט 235 מיליאן דאלאר ביי א שאצונג פון 4.5 ביליאן, און CoreWeave האט באקומען 2.3 ביליאן דאלאר הלוואות מיט Nvidia טשיפס אלס בטחון. גרויסע טעק-פירמעס ווי Google, Microsoft, Amazon און Apple האבן אריינגעשטעלט AI אין זייערע פראדוקטן. רעגירונגען האבן אויך אנגעהויבן צו ארבעטן: די ווייסע הויז האט לאנסירט די "AI Cyber Challenge", און כינע האט אנגעהויבן אויספירן נייע תקנות פאר AI.

מאדעלן און נייע לאנסירונגען

- **1 אויגוסט:** LLaMA-2 מיט 32K קאנטעקסט – Together AI האט ארויסגעלאזט LLaMA-2-7B-32K. דער מאדעל ניצט Position Interpolation און סיסטעם-אפטימיזאציעס פאר לאנגע אויפגאבעס, למשל צו פארשטיין עטליכע דאקומענטן אויף איינמאל, צו מאכן סיכומים און צו ענטפערן אויף פראגעס. ער ארבעט שנעלער סיי ביים ניצן און סיי ביים צוגעפאסטן טריינירן, און קומט מיט א דערנייערטן סאפטוועיר-סטעק.
- **3 אויגוסט:** Meta's AudioCraft – Meta האט ארויסגעגעבן AudioCraft, איין קאוד-באזע פאר אלע געדאנקען פון גענערירן קלאנג: מוזיק, קלאנג-עפעקטן און קאמפרעסיע.
- **4 אויגוסט:** IBM און NASA's געאגראפישער מאדעל – IBM און NASA האבן ארויפגעשטעלט אויף Hugging Face אן אפן-סאורס יסוד-מאדעל פאר געאגראפישע דאטא, געבויט אויף סאטעליט-בילדער פון NASA. מען באשרייבט אים אלס דער גרעסטער פון זיין סארט. ער האט פארבעסערט די רעזולטאטן מיט 15% ביי פאראויסזאגן מבול, און ער קען אויך סארטירן תבואה-פעלדער. דער ציל איז צו מאכן AI צוגענגלעך פאר קלימאט-וויסנשאפט.
- **7 אויגוסט:** Alibaba עפנט Tongyi Qianwen – Alibaba עפנט איר גרויסן שפראך-מאדעל Tongyi Qianwen פאר דרויסנדיגע אנטוויקלער.
- **9 אויגוסט:** StableCode פון Stability AI – Stability AI האט ארויסגעלאזט די ערשטע עפנטלעכע ווערסיע פון StableCode, אן אפענער שפראך-מאדעל וואס שרייבט קאוד.
- **10 אויגוסט:** נייער Claude Instant – Anthropic האט ארויסגעלאזט א דערנייערטע ווערסיע פון Claude Instant, איר שנעלערער און ביליגערער מאדעל. ער גיט לענגערע און בעסערע אויסגעשטעלטע ענטפערס און פאלגט בעסער אנווייזונגען וועגן פארמאט. ער איז אויך בעסער אין ארויסנעמען ציטאטן, אין פארשידענע שפראכן און אין ענטפערן אויף פראגעס. מען קען אים ניצן דורך אן API.

• **11 אויגוסט: Custom Instructions פאר יעדן** – OpenAI האט צוגעשטעלט "custom instructions" פאר אלע באניצער פון ChatGPT, אויך די וואס ניצן די אומזיסטע ווערסיע. דאס לאזט דעם באניצער באשטימען ווי ChatGPT זאל ענטפערן.

• **14 אויגוסט: WizardMath** – די WizardLM גרופע פון Microsoft האט ארויסגעלאזט וואס זי רופט דעם בעסטן אפן-סאורס מאדעל פאר מאטעמאטיק. ער איז שטארק ביי פראבלעמען וואס פאדערן עטליכע טריט, און ער גיט ערקלערונגען, כאטש ער דערפינדט אמאל זאכן. אין א פארשווג-ארבעט באשרייבט די גרופע ווי זי האט פארבעסערט דעם "reinforcement" חלק פון איר אויספיר-שטייגער. די רעזולטאטן שלאגן אסאך פון די בעסטע פארמאכטע מאדעלן אין מאטעמאטיק און לאגיש טראכטן.

• **14 אויגוסט: Platypus און Nvidia's אסיסטענט** – Orca און Platypus זענען צוזאמענגעשטעלט געווארן אין א שטארקן מאדעל אויף דער באזע פון Llama 2, וואס טאר נישט גענוצט ווערן פאר קאמערציעלע צוועקן. Nvidia מאכט גרינג צו טריינירן אן אייגענעם אסיסטענט פאר שפראך און בילדער, אויפגעבויט אויף איר Nemo מאדעל.

• **23 אויגוסט: פיין-טונינג פאר gpt-3.5-turbo** – OpenAI ערלויבט איצט צו טריינירן gpt-3.5-turbo אויף אייגענע דאטא. דער מאדעל וואס קומט ארויס געהערט צום באניצער, און די דאטא ווערט נישט גענוצט צו טריינירן אנדערע מאדעלן. דאס קען שטארק אראפברענגען די קאסטן, געבן מער קאנטראל איבער דעם מאדעל און פארבעסערן די רעזולטאטן. פיין-טונינג פאר GPT-4 קומט אין הארבסט. א נייע געצייג לאזט אויספירן דעם גאנצן פראצעס פון קאמאנד-ליין מיט טעקעס פון דער דיסק.

• **23 אויגוסט: SeamlessM4T פון Meta** – Meta האט ארויסגעלאזט SeamlessM4T, דער ערשטער איבערזעצונג-מאדעל וואס ארבעט סיי מיט פילע שפראכן און סיי מיט פארשידענע סארטן אינפארמאציע. ער נעמט אריין כמעט 100 שפראכן און גיט ארויס 35.

• **23 אויגוסט: IDEFICS פון Hugging Face** – Hugging Face האט ארויסגעלאזט IDEFICS, אן אפענע נאכמאכונג פון Flamingo. דאס איז א מאדעל וואס פארמישט טעקסט מיט בילדער, און קען ענטפערן פראגעס וועגן בילדער און טראכטן איבער זיי. ער איז געבויט אויף Llama v1 און קומט אין צוויי גרייסן: 8B און 70B.

• **24 אויגוסט: Midjourney's "(Region) Vary"** – Midjourney האט צוגעלייגט "(Region) Vary", א מעגלעכקייט צו טוישן א געוויסן טייל פון א בילד אינעווייניג (inpainting).

• **25 אויגוסט: Code Llama און זיינע נאכפאלגער** – Meta האט לאנסירט Code Llama, א מאדעל פאר קאד געבויט אויף Llama 2, אין דריי גרייסן: 7B, 13B און 34B. ער האט באקומען 53.7 אויף HumanEval און 56.2 אויף MBPP. ער קען שרייבן קאד און געפינען טעותים, ער איז טרייניר געווארן צו פאלגן אנווייזונגען, און זיין קאנטעקסט דערגרייכט ביז 100 טויזנט טאקענס. ווייניג טעג שפעטער האט Phind ארויסגעלאזט א צוגעפאסטע ווערסיע וואס שלאגט די מערץ-ווערסיע פון GPT-4 אויף א קאד-בענטשמארק. TLDR באמערקט אבער אז די היינטיגע GPT-4 איז נאך אלץ בעסער ווי יעדער אפן-סאורס מאדעל אין קאד. WizardLM האט נאכדעם מיט איר Evol instruct מעטאד אויך איבערגעיאגט TLDR. Phind באמערקט אבער אז מען דארף זיך נישט צו פיל פארלאזן אויף די רעזולטאטן, ווייל די אנווייזונגען זענען אויסגעשטעלט געווארן אזוי אז זיי זאלן גלייך העכערן די בענטשמארק-צאלן.

• **28 אויגוסט: Meta's PUG דאטא-סעט** – Meta AI האט ארויסגעלאזט PUG, דאטא-סעט מיט בילדער וואס זעען אויס ווי אמת'ע פאטאגראפיעס, געמאכט מיטן Unreal Engine. עס זענען דא דריי סארטן:

PUG: Animals פאר פארשונג אויף יסוד-מאדעלן, PUG: ImageNet צו פרובירן ווי פעסט בילד-סארטירער האלטן, און PUG: SPAR צו אפשאצן מאדעלן פון בילד מיט שפראך.

• **29 אויגוסט: ChatGPT Enterprise – OpenAI** האט לאנסירט ChatGPT Enterprise, א זיכערע און פריוואטע ווערסיע פאר פירמעס. זי איז ארום צוויי מאל אזוי שנעל, האט א קאנטעקסט פון 32K, קיין גרעניץ ווי פיל מען ניצט זי, און SOC2 זיכערהייט-באשטעטיגונג. לויט OpenAI האבן ארבעטער ביי 80% פון די Fortune 500 פירמעס זיך שוין איינגעשריבן צו ניצן ChatGPT.

• **29 אויגוסט: Qwen-VL פון Alibaba – Alibaba** האט לאנסירט צוויי נייע מאדעלן: Qwen-VL און Qwen-VL-Chat. לויט דער פירמע פארשטייען זיי בילדער און קענען פירן מער קאמפליצירטע שמועסן.

• **30 אויגוסט: Google's SynthID – Google** האט לאנסירט SynthID, א סימן וואס ווערט אריינגעלייגט אין א בילד. דער מענטש זעט אים נישט, אבער א ספעציעלע געצייג קען אים גרינג אנכאפן, און אזוי וועט מען וויסן אז א בילד איז געמאכט געווארן מיט AI.

• **31 אויגוסט: Duet AI אין Google Workspace** – Google שטעלט צו Duet AI אין אלע איר Workspace אפליקאציעס, אריינגערעכנט Gmail, Drive, Slides, און Docs, פאר 30 דאלאר א חודש. מער ווי א מיליאן מענטשן האבן אים שוין אויספרובירט.

• **31 אויגוסט: Palmyra פון Writer – Writer** האט ארויסגעלאזט Palmyra, א מאדעל וואס איז נישט געבויט אויף Llama. ער איז טרייניר געווארן אויף 800 ביליאן טאקעס און 70 טויזנט קערפול אויסגעקליבענע אנווייזונגען. ער איז שטארק אין פאלגן קאמפליצירטע אנווייזונגען, און מען קען אים גרינג צופאסן מיט ווייטערדיגן טריינירן.

פארשונג און טעכניק

• **1 אויגוסט: פארשידענע פארשונג-ארבעטן – TrafficSafetyGPT** איז א שפראך-מאדעל וואס איז צוגעפאסט געווארן מיט ספעציעלע דאטא פאר זיכערהייט אויף די וועגן. אן אנדערע ארבעט שלאגט פאר צוויי אלגאריטמען וואס טוישן אויטאמאטיש ווי שווער יעדע אויפגאבע וועגט ביים לערנען עטליכע אויפגאבעס אויף איינמאל. א דריטע ארבעט בייט א שטייגער צו שאפן ריזיגע דאטא-סעטס פאר טריינירן AI אגענטן וואס נאוויגירן לויט וואס זיי זעען.

• **1 אויגוסט: ווען GPT-3.5 איז ביליגער ווי Llama 2** – א בלאג פון Cursor ווייזט אז טריינירן, פראצעסירן דעם פראמפט און גענערירן טעקסט קאסטן יעדעס באזונדער, און ווערן אפגעהאלטן דורך פארשידענע רעסורסן פון קאמפיוטער. דערפאר קאסט געוויסע ארבעט מער אויף Llama 70B ווי צו באצאלן פאר GPT-3.5.

• **2 אויגוסט: געצייג-באנוץ, דיסטילאציע, מוזיק-באשרייבונג – ToolLlama** ווייזט אז מען קען זאמלען מער ווי 40 טויזנט APIs און טריינירן אן אפענעם מאדעל צו ניצן זיי. א נייער "knowledge distillation" שטייגער העלפט געוואלטע געצן נישט צו פארגעסן פריערדיגע לעקציעס ווען זיי לערנען נייע. פארשער האבן געזאמלט 2.2 מיליאן באשרייבונגען פון מוזיק, און דער מאדעל וואס זיי האבן טרייניר געווארן דערויף האט געוויזן זייער גוטע רעזולטאטן.

• **2 אויגוסט: טאנץ-וידעאס און 3D סצענעס – DisCo** שאפט ווידעאס פון מענטשן וואס טאנצן, וואס זעען אויס ווי אמת'ע. ער האלט אפ ווי דער מענטש און דער הינטערגרונט זעען אויס, און ער פאסט זיך צו אויך

צו נייע מענטשן און פאזעס. Symphonies פילט אויס די לעכער אין 3D סצענעס פאר אויטאמאטישע קארס.

• **3 אויגוסט: ראיה-מאדעלן און LISA** – א נייע נעץ פארבעסערט ווי גענוי מען סארטירט מוח-געשוויילסטן אין MRI בילדער. YOLOBench מעסט מער ווי 550 YOLO מאדעלן אויף פיר דאטא-סעטס און פארשידענע הארדוועיר. "Focused Linear Attention" מאכט Vision Transformers מער עפעקטיוו. LISA צייכנט אן די פלעקער אין א בילד וואס פאסן צו א קאמפליצירטע אדער אומקלארע טעקסט-אנווייזונג, אן אויפגאבע וואס הייסט "reasoning segmentation".

• **3 אויגוסט: MIT's Liquid Neural Networks – MIT CSAIL's "Liquid Neural Networks"** זענען קליינע מאדעלן וואס זענען שטארק אין ראבאטיק און אויטאמאטישע קארס. זיי פאסן זיך צו, פאדערן ווייניגער קאמפיוטער-כח, און זענען בעסער ווי געווענליכע מאדעלן אין א סביבה וואס טוישט זיך. מיט פעסטע דאטא-באזעס האבן זיי אבער שוועריגקייטן.

• **4 אויגוסט: SoftMoE און זעלבסט-קאנטראל – SoftMoE** פארבעסערט די ספארס MoE מאדעלן: יעדער "עקספערט" באקומט א טייל פון די טאקענס דורך א ווייכע צוטיילונג. ער איז אזוי שטארק ווי ViT-Huge/14, אבער קאסט אסאך ווייניגער צו ניצן. אן אנדער ארבעט ווייזט ווי שפראך-מאדעלן קענען אליין אנכאפן זייערע טעותים אין פראבלעמען פון עטליכע טריט. א דריטע ארבעט צייכנט אפ אביעקטן אין בילדער אן טריינירן.

• **7 אויגוסט: הערן וואס מען טיפט** – אויב מען טיפט ווען דער מיקראפאן איז נישט אפגעשטעלט, קען א נעוראלע נעץ אויסרעכענען וואס מען טיפט מיט 93% גענויקייט. דאס איז אסאך בעסער ווי אלע פריערדיגע מעטאדן.

• **7 אויגוסט: PerceptionCLIP, CoThought, Magic123** – PerceptionCLIP פארבעסערט ווי CLIP סארטירט בילדער: ער זוכט קודם אויס די אייגנשאפטן פונעם הינטערגרונט. CoThought ניצט GPT-3.5-turbo צו איבערארבעטן דאטא, כדי צו טריינירן קליינע "בעיבי" מאדעלן. Magic123 שאפט 3D אביעקטן פון איין 2D בילד, און איז געשפרונגען פאראויס אין קוואליטעט.

• **8 אויגוסט: אפשאצן מאדעלן** – א נייער דאטא-סעט פרובירט צי AI זעט בילדער אזוי ווי א מענטש, מיט דריי גרופעס: Must-Act, Must-Abstain און Uncertain. אן אנדער ארבעט אפשאצט שפראך-מאדעלן דורך נעצן וואס נאכמאכן ווי מען רעצענזירט אקאדעמישע ארבעטן. MM-Vet פרובירט מאדעלן אויף אויפגאבעס מיט בילדער און ווערטער צוזאמען.

• **9 אויגוסט: AgentBench און Anthropic's פארשונג** – AgentBench פרובירט 25 שפראך-מאדעלן אלס אגענטן, און קאמערציעלע מאדעלן זענען געווען בעסער ווי אפן-סאורס. Anthropic האט מיט "influence functions" אויסגעפארשט ווי טריינירן-ביישפילן ווירקן אויפן ענדגילטיגן מאדעל. זי האט געפונען אז די השפעה פאלט אראפ צו נול ווען מען דרייט איבער די סדר פון וויכטיגע פראזעס. DAT פארבעסערט ווי מען מאכט בילדער שארפער.

• **10 אויגוסט: Grokking: אויסווענדדיג לערנען אדער פארשטיין** – Google's PAIR גרופע האט פארעפנטלעכט ערקלערונגען וועגן "Grokking", א טויש ביים טריינירן וואס ווייזט מעגלעך אז דער מאדעל גייט איבער פון אויסווענדדיג לערנען צו פארשטיין. זיי האבן אויך צוריק-אויסגעפארשט די לייזונג וואס א קליינער מאדעל געפינט.

- **10 אויגוסט:** סינטעטישע דאטא קעגן חנופה – מאדעלן נעמען אן די מיינונג פונעם באניצער, און דאס געשעט מער ביי גרעסערע מאדעלן און ביי מאדעלן וואס זענען טרייניר געווארן צו פאלגן אנווייזונגען. מען קען דאס פארמיידן מיט פשוט'ע סינטעטישע דאטא, אן צו שאטן די אלגעמיינע רעזולטאטן.
- **11 אויגוסט:** In-context לערנען מיט Llama-2 – ווען מען זוכט ארויס 9 ביישפילן און לייגט זיי אריין אין פראמפט, באקומט Llama-2 א 7 מאל העכערן זיג-ראטע קעגן text-davinci-003, אן צו טוישן די ווייטס. דאס ברענגט אים צו דער מדרגה פון מאדעלן וואס זענען טרייניר געווארן מיט alignment. אנדערע ארבעטן פארבעסערן 3D דעטעקציע און ווי מען פארבינדט בילדער מיט טעקסט.
- **14 אויגוסט:** Follow Anything און זומען – "Follow Anything" איז א ראבאט וואס דערקענט און גייט נאך יעדן אביעקט אין רעאל-צייט, לויט א טעקסט, א בילד אדער א קליק, מיט נאר א געווענליכן לעפטאפ. ווען מען זומט אריין אויף וויכטיגע טיילן פון א בילד פארן סארטירן, באקומט מען 98% אויף ImageNet.
- **15 אויגוסט:** ראיא און Lidar – EMMS שאצט אפ ווי גוט א מאדעל וועט ארבעטן אויף פארשידענע אויפגאבעס אן אים צו טריינירן. א לייכטע RGB-D נעץ איז געמאכט פאר ראבאטן. MS3D++ פאסט צו 3D דעטעקטארן צו נייע סארטן Lidar. DatasetDM שאפט אן א סוף סינטעטישע בילדער מיט באצייכענונגען.
- **16 אויגוסט:** קאוד-מאדעלן און Bayesian Flow Networks – א נייע ארבעט טריינירט קאוד-מאדעלן אן ניצן דאטא פון GPT-4, און דערגרייכט 46.7 אויף HumanEval. TLDR באמערקט אבער אז דאס ווייזט ווי גרויס דער אונטערשייד איז צווישן פארמאכטע און אפענע מאדעלן. DCI-VTON ניצט דיפוזיע פאר ווירטועל אנפרובירן קליידער. Bayesian Flow Networks זענען א נייע קלאס פון גענעראטיווע מאדעלן פאר סיי דיסקרעטע און סיי קאנטיניוירליכע דאטא.
- **17 אויגוסט:** צען אפענע פראבלעמען – Chip Huyen רעכנט אויס 10 אפענע פראבלעמען אין דער פארשונג פון שפראך-מאדעלן: וואס דער מאדעל דערפינדט זאכן, לענג פונעם קאנטעקסט, פארשידענע סארטן אינפארמאציע, נייע ארכיטעקטורן, אלטערנאטיוון צו GPUs, און ווייטער.
- **17 אויגוסט:** פארשידענע מעטאדן – L3DMC פארבעסערט לערנען אן אויפהער. Solo Performance Prompting (SPP) ניצט עטליכע "פערזאנען" אין איין מאדעל צו לייען פראבלעמען. CommitChronicle איז א נייער דאטא-סעט צו פארענדיגן commit מעלדונגען מיט דער היסטאריע פון פריערדיגע commits.
- **18 אויגוסט:** MyoSuite פון Meta AI – MyoSuite סימולירט די מוסקלען און ביינער פונעם מענטשליכן קערפער, און טריינירט אגענטן צו באוועגן הענט און פיס מיט מענטשליכער געשיקטקייט.
- **18 אויגוסט:** אויגן-בילדער און קערנלס – DDPM מיטן דאטא-סעט ReTree מאכט גרינגער צו שאפן בילדער פון דער רעטינע. DeDoDe געפינט בעסער די שליסל-פונקטן פאר 3D. KernelWarehouse מאכט דינאמישע קאנוואלוציע מער עפעקטיוו מיט ווייניגער פאראמעטערס.
- **21 אויגוסט:** SigLIP און 3D – א פארשונג-ארבעט ווייזט אז א פשוט'ערער שטייגער צו טריינירן, מיט א sigmoid loss, איז בעסער ווי GPA-3D. CLIP פארבעסערט 3D דעטעקציע אין נייע ערטער. א נייע מעטאד שאצט עפעקטיוו אפ בילדער וואס זענען געמאכט פון טעקסט.
- **22 אויגוסט:** AlphaZero קרעאטיוו – DeepMind האט געמאכט AlphaZero מער קרעאטיוו אין שאך: זיינע צוגן זענען מער פארשידענארטיג, און ער לייזט שווערע רעטענישן. StableVideo רעדאגירט ווידעאס און האלט די זעלבע אויסזען פון אביעקטן. א נייער אלגאריטם פארבעסערט טשעטבאטס וואס שפילן אויס פיקטיווע פערזאנען.

- **23 אויגוסט:** קליינע אביעקטן און NeRF – CFNet דערקענט בעסער קליינע אביעקטן. DReg-NeRF פארבינדט עטליכע NeRF סצענעס אן מענטשליכע הילף.
- **24 אויגוסט:** זעלבסט-טריינירן, 3D, נאכגיין – א סקירה איבער semi- און self-supervised לערנען. א נייע מעטאד זוכט בילדער ענליך צו א געגעבן בילד, מיט ענדערונגען וואס א באשרייבונג זאגט. IT3D מאכט טעקסט-צו-3D מער רעאליסטיש. TOPIC גייט נאך עטליכע אביעקטן אין ווידעאס, מיט א נייעם דאטא-סעט BEE23.
- **25 אויגוסט:** נייע בענטשמארקס – Stanford האט געבויט א גרויסן בענטשמארק פאר לאגיש טראכטן אין געזעץ. Google האט געבויט א בענטשמארק פאר דערציילן מעשיות דורך גענערירטע ווידעאס. MTM-Tracker גייט נאך 3D אביעקטן מיט Lidar. Cheetah איז א מאדעל וואס פארשטייט פארמישטע אנווייזונגען פון בילד און טעקסט, אפגעשאצט מיטן נייעם I4 בענטשמארק.
- **28 אויגוסט:** CAPTCHA'ס זענען שוין אומניצליך – א פארשונג ווייזט אז AI באטס לייזן CAPTCHA'ס שנעלער און גענויער ווי מענטשן.
- **28 אויגוסט:** CALM, Apple'ס FineRecon – CALM איז א טעסט וואס ווייזט אז געוויסע גרויסע מאדעלן זענען מער פארדרייט ווי קלענערע. Apple'ס FineRecon ווייזט אויסערגעווענליכע רעזולטאטן אין אויפבויען 3D אביעקטן. MGDN פארמישט עטליכע בילדער אין איין קלאר בילד, און PViC דערקענט בעסער ווי מענטשן האנדלען מיט אביעקטן.
- **29 אויגוסט:** ליכט, קאמערעס, פנימער – א נייע מעטאד בויט אויף ווי ליכט פאלט אויף 3D אביעקטן פון ווייניג בילדער. א גראף-מאדעל גייט נאך אביעקטן איבער עטליכע קאמערעס. DFER-CLIP דערקענט ווי פנים-אויסדרוקן טוישן זיך.
- **30 אויגוסט:** UniPT, 3D Gaussian, DiffSmooth – UniPT מאכט אז transfer learning זאל ניצן ווייניגער זכרון. א נייע מעטאד מיט 3D Gaussians בויט אויף א פארעם פון א ווידעא פון א טעלעפאן אין עטליכע סעקונדעס אויף א CPU. DiffSmooth ניצט דיפוזיע צו מאכן מאדעלן פעסטער קעגן אטאקעס.
- **31 אויגוסט:** Algorithm of Thoughts – Microsoft האט פארגעשטעלט "Algorithm of Thoughts" (AoT), א שטייגער צו מאכן אז שפראך-מאדעלן זאלן טראכטן עפעקטיווער און מער ווי א מענטש.
- **31 אויגוסט:** מעדיצינישע סעגמענטאציע און נאכגיין – MIT'ס UniverSeg איז א מאדעל אויפגעבויט אויף UNet, וואס ארבעט אויף פארשידענע מעדיצינישע בילדער. Meta'ס CoTracker גייט נאך פיקסלען אין ווידעאס, אפילו ווען עפעס פארדעקט זיי אדער זיי באוועגן זיך שנעל. DySample העכערט די רעזאלוציע פון בילדער אויף א לייכטן אופן. א נייע מעטאד דערקענט 3D פאזעס פונעם קערפער לויט סיי ארט און סיי צייט.

געצייג פאר פראגראמירער און אפן-סאורס

- **1 אויגוסט:** Ollama, FacTool, Rust Llama – Ollama, לויפט, שאפט און טיילט שפראך-מאדעלן. FacTool געפינט פאקטישע טעותים אין טעקסט וואס מאדעלן שרייבן. Rust Llama איז א נאכמאכונג פון llama2.c אין Rust.

- **4 אויגוסט:** מוסטערן פאר בויען מיט LLMs – Eugene Yan באשרייבט זיבן פראקטישע מוסטערן ווי צו בויען סיסטעמען און פראדוקטן מיט שפראך-מאדעלן.
- **7 אויגוסט:** Functionary – Functionary איז א מאדעל וואס פארשטייט פונקציעס און פירט זיי אויס. ער ארבעט מיט JSON Schema, ענליך צו OpenAI's function calling.
- **8 אויגוסט:** MiniChain, LanceDB און מער – MiniChain לאזט פארבינדן עטליכע אויפרופן פון מאדעלן איינע נאך די אנדערע מיט דעקאראטארן און YAML, און מען קען שרייבן א טשעטבאט אין בערך 20 שורות. LanceDB 2.0 איז א סערווער-לאזע וועקטאר דאטא-באזע. Hugging Face'ס 'generate()' שטיצט איצט WebGPU-Torch classifier free guidance. TorchScale איז א ביבליאטעק פון Microsoft צו פארגרעסערן Transformers.
- **9 אויגוסט:** טריינירן מיט איין קליק אויף Hugging Face – Hugging Face ארבעט צוזאמען מיט AWS און Nvidia, כדי מען זאל קענען טריינירן מאדעלן מיט איין קליק, נאר דורך ארויפלאדן די דאטא. אנדערע געצייג: א קליינע ווערסיע פון SDXL אין עטליכע הונדערט שורות; Humanscript; Sweep, אן אפן-סאורס AI "יונגער פראגראמירער" וואס מאכט פון GitHub issues א pull request; א לערנונג ווי צו ניצן Swift אויף Apple מכשירים; און Agentflow.
- **9 אויגוסט:** Project IDX – Google האט לאנסירט Project IDX, א פלאץ אין בראווער צו אנטוויקלען אפליקאציעס, מיט AI.
- **10 אויגוסט:** Magentic און Llama לערנונגען – SimplyRetrieve לייגט אריין פריוואטע דאטא אין א שפראך-מאדעל אן טריינירן. Magentic פארבינדט LLMs מיט Python קאוד. עס איז דא א לערנונג ווי צו בויען Llama פון אנהייב, א Colab ווי צו צופאסן Llama 2 מיט QLoRA, און api2ai.
- **11 אויגוסט:** Stanford'ס דארף-אגענטן אפן – Stanford האט געעפנט דעם קאוד פון איר סימולירטן דארף מיט מענטשן וואס לעבן דארט. Chidori איז א פלאץ צו בויען AI אגענטן. Llama-2-Onnx איז אן אפטימיזירטע ווערסיע פון Llama 2.
- **14 אויגוסט:** Supabase, ControlNet, Council – Supabase Studio 3.0 ברענגט AI אינעם SQL עדיטאר. עס איז דא א ControlNet פאר MosaicML SDXL. האט א זיכערן פלאץ צו פרובירן גענערירטן קאוד. Council'ס ChainML איז א פלאץ פאר אגענטן וואס ארבעטן צוזאמען. Mirage האט געעפנט איר טעקסט-צו-3D. Preemo האט גענומען Hugging Face'ס text-generation-inference און אנגעהויבן אנטוויקלען זיין אייגענע אפן-סאורס ווערסיע, נאכדעם וואס Hugging Face האט צוגענומען די קאמערציעלע ליצענץ.
- **15 אויגוסט:** Fooocus און GPU פירער – Fooocus שאפט בילדער אפליין, אומזיסט, און מען דארף נישט צופאסן קיין סעטינגס. א פירער ערקלערט וועלכע וואלקן-פירמע מען זאל אויסקלייבן פאר GPUs.
- **16 אויגוסט:** FoodSAM, DoctorGPT, OpenDream – FoodSAM צייכנט אפ עסן אין בילדער. DoctorGPT איז אן אפן-סאורס מאדעל וואס דורכפאלט נישט דעם אמעריקאנער דאקטער-עקזאמען (USMLE), ארבעט אפליין און האלט פריוואט די געזונט-דאטא. OpenDream ברענגט שיכטן און רעדאגירן וואס קען צוריקגעדרייט ווערן צו Stable Diffusion.
- **17 אויגוסט:** Pixi – Pixi איז א נייער פעקל-פארוואלטער פאר Conda, געשריבן אין Rust.

- **18 אויגוסט: Rift, DeepEval, LoRA** – Rift איז אפן-סאורס אינפראסטרוקטור פאר א פלאץ צו אנטוויקלען קאוד וואס איז געבויט אויף DeepEval AI. אפשאצט LLM סיסטעמען אין Python. א נייער פלאץ ווייזט LoRAs פאר SDXL וואס די קהילה האט טרייניר געווארן. עס זענען אויך דא ביישפילן פון LangChain אין דער פראקטיק.
- **21 אויגוסט: 3 Dolma טריליאן טאקענס** – Allen Institute for AI האט ארויסגעלאזט Dolma, אן אפענער דאטא-סעט פון 3 טריליאן טאקענס, כדי פארשער זאלן קענען שטודירן ווי דאטא ווירקט אויף מאדעלן ביי גרויסע מאסן. אויך: די סביבה פונעם ראבאט-פיאניסט איז געעפנט געווארן; עס איז דא אן אומאפיציעלער קאוד פאר Bayesian Flow Networks; און txtai.
- **22 אויגוסט: Hugging Face's Rust פריימווערק** – Hugging Face האט ארויסגעלאזט א קליין און זייער שנעל ML פריימווערק אין OpenPipe. Rust. פארגלייכט פראמפטס און איבערזעצט זיי פון איין מאדעל צום אנדערן. RLIPv2 פארבינדט בילדער מיט ווערטער שנעלער. FlexFlow איז א פריימווערק צו לויפן מאדעלן אין פראדוקציע.
- **23 אויגוסט: Whisper API און SynJax** – Whisper API איז א טראנסקריפציע-API וואס מען קען אליין האלטן אויף אן אייגענעם סערווער. SynJax פון DeepMind געבט געצייג פאר מעגלעכקייטן איבער סטרוקטורן ווי ביימער און סעקווענצן.
- **24 אויגוסט: OneDiffusion און GGML** – OneDiffusion מאכט גרינגער צו לויפן און צופאסן דיפוזיע-מאדעלן. Llama.cpp פרובירט א נייעם פארמאט וואס זאל שטיצן מער מאדעלן.
- **25 אויגוסט: Prompt2Model און fp8** – Prompt2Model שאפט א קליינעם, עפעקטיוון מאדעל פון א באשרייבונג אין פשוט'ע ווערטער, און ער קען זיין בעסער ווי געוויסע גרויסע מאדעלן. Unit's Graphcore Scaling מאכט גרינג און סטאביל צו טריינירן אין fp8. עס איז אויך דא א לערנונג ווי צו ניצן DPO אויף Llama 2 אן אן reward מאדעל, און אן AutoML וועב-אפליקאציע.
- **28 אויגוסט: Apple's שנעלער ViT, DeepSpeed** – Apple האט א vision transformer וואס לויפט אין ווייניגער ווי 1 מילי-סעקונדע. DeepSpeed פון Microsoft האט צוגעלייגט sequence parallelism פאר לענגערע קאנטעקסטן. OpenLemur פרובירט צו האלטן קאוד-מאדעלן ביי גוטע שפראך-פעאייגקייטן. עס זענען אויך דא OpenCoPilot און א פאבריק פון קאמיקס.
- **29 אויגוסט: קלאנג, Lagent, GPT Pilot** – אן אפן-סאורס מאדעל שאפט שנעל ביטס, קלאנג-עפעקטן און פשוט'ע שמועסן. Lagent איז א לייכט פריימווערק פאר אגענטן. GPT Pilot שרייבט גאנצע אפליקאציעס, און דער פראגראמירער גייט נאך די ארבעט.
- **30 אויגוסט: LangFuse און DSPy** – LangFuse איז אן אפן-סאורס געצייג צו אנקוקן און אנאליזירן LLM אפליקאציעס. עס איז דא א ליסטע פון לייענען וועגן "adaptive computation". DSPy פאראייניגט פראמפטינג און טריינירן אין Python מאדולן.
- **31 אויגוסט: BatteryML און CUDA graphs** – BatteryML איז א געצייג פאר פארשונג וועגן ווי באטעריעס ווערן שוואכער. מיט torch.compile און CUDA graphs איז Llama 2 געווארן 2.3 מאל שנעלער. Phind איז א זוכמאשין מיט AI און א שותף פאר פראגראמירן.

פירמעס, געלט און ביזנעס

- **1 אויגוסט: Dell און Intel** – Dell לאנסירט "Dell Generative AI Solutions", הארדוועיר און באדינונגען, צוזאמען מיט Nvidia. Intel CEO Pat Gelsinger האט געזאגט אז יעדער צוקונפטיגער פראדוקט וועט האבן AI, אנגעהויבן מיטן Meteor Lake טשיפ, און אז AI דארף לויפן אויפן אייגענעם מכשיר אנשטאט אין וואלקן.
- **1 אויגוסט: Amazon'ס נייע AI גרופע** – Amazon CEO Andy Jassy פירט אליין א נייע גרופע וואס ארבעט אויף די פירמע'ס "מערסט אמביציעזע" AI מאדעלן.
- **2 אויגוסט: טשעטבאטס אין אפליקאציעס** – Uber אנטוויקלט א טשעטבאט פאר איר אפליקאציע. Meta וועט מעגלעך לאנסירן AI "פערזאנען" אין סעפטעמבער. LinkedIn פלאנט א AI "קאוטש" פאר זוכן ארבעט. YouTube פרובירט סיכומים פון ווידעאס וואס AI שרייבט.
- **3 אויגוסט: White Castle און Kickstarter** – White Castle וועט אריינשטעלן AI קול-אסיסטענץ פון SoundHound אין מער ווי 100 drive-thrus ביז 2024, און זאגט אז מען וועט נישט אפזאגן קיין ארבעטער. Kickstarter פאדערט אז פראיעקטן וואס ניצן AI זאלן אפענטלעך ערקלערן ווי זיי ניצן עס און פון וואו די דאטא קומט.
- **4 אויגוסט: CoreWeave'ס 2.3 ביליאן** – CoreWeave האט אויפגענומען 2.3 ביליאן דאלאר הלוואות מיט Nvidia טשיפס אלס בטחון, כדי צו בויען AI דאטא צענטערס, און אזוי פארמיידט זי צו פארקויפן אקציעס. די פירמע, וואס האט אנגעהויבן מיט קריפטא, ווערט געשאצט ארום 2 ביליאן און פארמעסט זיך מיט Microsoft, Google און Amazon. דער מארק פאר AI אינפראסטרוקטור ווערט געשאצט צו וואקסן ביז 422.55 ביליאן דאלאר אין 2029.
- **4 אויגוסט: GPT-5 טרעידמארק** – OpenAI האט אריינגעגעבן א בקשה פאר א טרעידמארק אויפן נאמען GPT-5.
- **7 אויגוסט: MatX, Google Assistant, Apple** – מענטשן וואס האבן געבויט TPUs האבן אויפגעשטעלט MatX, א נייע פירמע פאר א טשיפ ספעציעל פאר Transformers. Google וועט ברענגען גענעראטיוו AI אין Assistant, און האט אפגעזאגט א טייל פון דער גרופע וואס האט געארבעט דערויף. Tim Cook זאגט אז Apple אינוועסטירט שוין "יארן" אין גענעראטיוו AI.
- **8 אויגוסט: Salesforce און Bing Chat** – Einstein Studio'ס Salesforce לאזט פירמעס אריינברענגען זייערע אייגענע מאדעלן, אנגעהויבן מיט Microsoft SageMaker. Amazon'ס Bing Chat קומט אין אנדערע בראווערס, און שוין אין Chrome.
- **9 אויגוסט: Glaive און גרויסע טעק** – Glaive האט אויפגענומען 3.5 מיליאן דאלאר אין א seed געלט-רונדע פאר קליינע מאדעלן געבויט אויף סינטעטישע דאטא, וואס זענען 2 ביז 20 מאל שנעלער און לכל הפחות 2 מאל ביליגער. Google און Universal Music רעדן וועגן א ליצענץ פאר מעלאדיעס און קולות פון זינגערס פאר מוזיק וואס AI שאפט. עס איז אויך דא א סקירה ווי גרויסע טעק פירמעס האבן גערעדט וועגן AI אין זייערע פערטל-יארעס באריכטן.
- **10 אויגוסט: Inworld AI** – Inworld AI האט אויפגענומען מער ווי 50 מיליאן דאלאר, אנגעפירט דורך Lightspeed, ביי א שאצונג פון מער ווי 500 מיליאן. אזוי איז זי געווארן דער סטארטאפ אין AI פאר שפילן וואס האט באקומען די מערסטע געלט.

- **11 אויגוסט:** **SoftBank, Pinecone, Meta** – SoftBank שטעלט אויף GreenBox Systems צוזאמען מיט Symbolic, מיט 100 מיליאן דאלאר, און וועט קויפן AI סיסטעמען פאר לאגערן פאר 7.5 ביליאן דאלאר. Pinecone ערקלערט ווי זי איז דערגאנגען צו 10 טויזנט איינשרייבונגען א טאג. Meta האט צעלאזט איר גרופע פאר protein folding, צוליב דעם דרוק צו פארקויפן AI פראדוקטן.
- **14 אויגוסט:** **Serve Robotics** גייט עפנטלעך – Serve Robotics, וואס איז ארויסגעוואקסן פון Postmates און ווערט געשטיצט דורך Uber און Nvidia, גייט עפנטלעך דורך א reverse merger.
- **15 אויגוסט:** **Amazon, Apple, Humane** – Amazon פירט איין סיכומים פון רעצענזיעס אויף פראדוקטן, געשריבן מיט AI. Apple ניצט אן אינערליכן טשעטבאט פאר אירע ארבעטער. Humane וועט דערציילן מער וועגן איר AI Pin אין אקטאבער.
- **15 אויגוסט:** **AI.com איבער צו x.AI** – AI.com ווייזט איצט צו Elon Musk'ס x.AI אנשטאט צו ChatGPT.
- **16 אויגוסט:** **Google Search און Kneron** – Google'ס Search Generative Experience באקומט ערקלערונגען פון ווערטער, הילף מיט פראגראמירן און סיכומים. Kneron וועט שיקן KL730 טשיפס ביזן סוף יאר, וואס זענען 3 ביז 4 מאל מער עפעקטיוו מיט ענערגיע.
- **17 אויגוסט:** **OpenAI קויפט Global Illumination** – OpenAI האט געקויפט Global Illumination, א גרופע וואס האט פריער געארבעט ביי Facebook, YouTube און Riot, כדי צו פארבעסערן ChatGPT. דאס קומט נאכדעם וואס עטליכע פראדוקט-פירער זענען אוועק פון OpenAI. DeepMind פון AI "לעבנס-קאוטש". FleetWorks, געגרינדעט דורך געוועזענע ארבעטער פון Uber Freight און Airbnb, ניצט AI צו פירן די טעלעפאן-רופן אין טראנספארט.
- **18 אויגוסט:** **Toyota/Pony.ai און Sakana** – Toyota און Pony.ai וועלן מאסנווייז פראדוצירן robotaxis אין כינע, מיט מער ווי 140 מיליאן דאלאר. Sakana, א נייע פארשונג-לאבאראטאריע אין טאקיא, געגרינדעט דורך געוועזענע פון Google Brain און Stability, זוכט נייע מאדעלן וואס זענען אינספירירט פון דער נאטור.
- **22 אויגוסט:** **Snap'ס Dreams** – Snap לאנסירט Dreams, וואס שטעלט די בילדער פונעם באניצער אריין אין פאנטאסטישע ערטער מיט AI.
- **23 אויגוסט:** **Python אין Excel** – Microsoft ברענגט Python אריין אין Excel, פאר אנאליז פון דאטא, טשארטס און רייניגן דאטא.
- **24 אויגוסט:** **Ideogram און Windows** – Ideogram געוועזענע פון Google Brain וואס האבן געארבעט אויף בילדער האבן געגרינדעט Ideogram. Microsoft פארבירט אריינצולייגן AI אין Photos, Snipping Tool און Paint.
- **25 אויגוסט:** **Hugging Face: 235 מיליאן** – Hugging Face האט אויפגענומען 235 מיליאן דאלאר אין א Series D געלט-רונדע, מיט Google, Amazon, Nvidia, Intel, AMD, Qualcomm, IBM, Salesforce און Sound Ventures, ביי א שאצונג פון 4.5 ביליאן דאלאר.
- **29 אויגוסט:** **Compute אינדעקס, DoorDash** – דער State of AI באריכט רעכנט ווי פיל A100 טשיפס יעדע פירמע פארמאגט, און באמערקט אז 7 פארשונג-ארבעטן האבן גענוצט Huawei'ס טשיפ. DoorDash לאנסירט א AI ענטפער-סערוויס פאר רעסטאראנען.

- **30 אויגוסט:** Poe – Amazon/Fig און Poe האט צוגעלייגט זוכן נאך באטס און א פלאן פאר פירמעס. Amazon האט געקויפט Fig, א סטארטאפ פון YC וואס האט אויפגענומען 2 מיליאן דאלאר.
- **31 אויגוסט:** OpenAI נאענט צו א ביליאן – OpenAI קען דערגרייכן א ביליאן דאלאר הכנסות אין די קומענדיגע 12 חדשים, אסאך שנעלער ווי איר אייגענע פאראויסזאגונג. Samsung האט לאנסירט Food, אן אפליקאציע פאר עסן און רעצעפטן מיט AI, אין 8 שפראכן און 104 לענדער.

רעגירונגען, געזעצן און זיכערהייט

- **2 אויגוסט:** כינע'ס תקנות און שווינדלער – א באריכט פון Matt Sheehan פון Carnegie Endowment ערקלערט ווי כינע מאכט אירע AI תקנות. גנבים אויפן אינטערנעט בויען טשעטבאטס ווי WormGPT און FraudGPT פאר phishing און שעדליכע סאפטוועיר.
- **3 אויגוסט:** Apple נעמט אראפ AI אפליקאציעס אין כינע – Apple האט אראפגענומען פילע AI אפליקאציעס פון איר App Store אין כינע, אריינגערעכנט OpenCat, פאר די נייע תקנות וואס פאדערן א ליצענץ. די תקנות קענען אפשרעקן קליינע אנטוויקלער.
- **4 אויגוסט:** וואס Congress באטראכט – Vox ערקלערט די AI כללים וואס Congress באטראכט: קאפירעכט, פריוואטקייט, און אלגאריטמען וואס זענען נישט ישר'דיג. ער ערקלערט אויך די פארשלאגן פאר נייע אויפזעצער-אינסטיטוציעס און געלט פאר פארשוונג. OpenAI און Microsoft שטיצן ליצענצן און קאנטראלן פאר שטארקע מאדעלן.
- **8 אויגוסט:** מען קען אפשטעלן OpenAI'ס קרויכער – OpenAI לאזט וועבזייטלעך אפשטעלן איר web crawler פון זאמלען דאטא פאר טריינירן GPT.
- **10 אויגוסט:** 82% ווילן פארלאנגזאמען – אין אן Axios אנקעטע פון 1001 מענטשן האבן 82% אמעריקאנער געזאגט אז מען זאל פארלאנגזאמען די אנטוויקלונג פון AI. 52% מיינען אז די רעגירונג דארף רעגולירן, און א גרויסע מערהייט גלייבט נישט אז די טעק ריזן וועלן זיך אליין קאנטראלירן.
- **10 אויגוסט:** AI Cyber Challenge – די ווייסע הויז האט לאנסירט ביי Black Hat א פארמעסט, געפירט דורך DARPA, צו בויען AI וואס געפינט און פאררעכט שוואכקייטן אין סאפטוועיר. עס זענען דא 20 מיליאן דאלאר פרייזן, Google און OpenAI זענען מיט אין דעם, און די געווינער ווערן ערמוטיגט צו עפענען זייערע לייוונגען.
- **11 אויגוסט:** צייטונגען פאדערן כללים – מעדיע-פירמעס האבן אונטערגעשריבן אן אפענעם בריוו וואס פאדערט אז מען זאל אפענטלעך ווייזן די טריינירן-דאטא, און אז מען דארף בעטן רשות פון די בעלי-זכות פאר מען ניצט זייער דאטא.
- **15 אויגוסט:** אויטאמאטישע טעקסיס אין San Francisco – די רעגולירער האבן ערלויבט Cruise און Waymo צו פירן מענטשן פאר געלט 24 שעה א טאג.
- **16 אויגוסט:** GPT-4 אלס מאדעראטאר – OpenAI ניצט GPT-4 צו אויספירן אירע כללים וועגן אינהאלט, און שלאגט פאר אז פלאטפארמעס זאלן אים ניצן צו שרייבן כללים און באשליסן, כדי צו נעמען די לאסט אראפ פון מענטשן. TLDR פרעגט צי AI זאל באשטימען ווער טאר זאגן וואס.

- **18 אויגוסט:** **New York Times** קעגן **OpenAI** – די **New York Times** גרייט זיך מעגלעך צו קלאגן OpenAI. דאס קען צווינגען צו פארניכטן ChatGPT'ס דאטא-סעט, מיט שטראפן ביז 150 טויזנט דאלאר פאר יעדן שטיק אינהאלט.
- **23 אויגוסט:** **זיכערן AI** – **Greylock** שרייבט אז מען דארף טראכטן ווי מען וועט מיסברויכן AI אין דער זעלבער צייט וואס זי וואקסט שנעל.
- **29 אויגוסט:** **ווי צו פירן AGI** – אן ארטיקל באטראכט ווי מען זאל פירן AGI: אפשאצונגען, קאנטראל ווער עס האט צוטריט, און אינטערנאציאנאלע אויפזיכט.
- **31 אויגוסט:** **קאפירעכט פאר AI קונסט** – דער געריכט אין DC האט באשלאסן אין **Thaler v. Perlmutter** אז קונסט וואס AI שאפט פאדערט א מענטשליכן מחבר כדי צו באקומען קאפירעכט. דער באשלוס איז געבויט אויף **Thaler**'ס טענה אז זיין AI האט געארבעט גאנץ אליין.

אנאליזן און מיינונגען

- **3 אויגוסט:** **אגענטן און alignment** – אן ארטיקל אויף **LessWrong** טענהט אז אגענטן ווי **Auto-GPT** און **Baby AGI** קענען שנעל ברענגען AGI. זיי ברענגען נייע שוועריגקייטן, אבער ווייל זיי טראכטן אויף ענגליש, איז גרינגער צו פארשטיין וואס זיי טוען.
- **4 אויגוסט:** **געזונט און 2023** – **a16z** מיינט אז AI וועט האבן די גרעסטע השפעה אין געזונט-וועזן, וואס איז 20% פון דער אמעריקאנער עקאנאמיע. **Not Boring** טענהט אז מען וועט געדענקען 2023 ווי דאס יאר ווען דער טעכנאלאגישער פארשריט האט ווידער אנגעהויבן לויפן.
- **7 אויגוסט:** **MoE, לאגיק, פארשריט** – אן ערקלערונג פון **Mixture of Experts**: ספארס מאדעלן ניצן נאר א טייל פון די פאראמעטערס פאר יעדן טאקען, און זענען עפעקטיווער און שטארקער. אן ארטיקל טענהט אז AI האט נאך עיקר'דיגע גרעניצן אין לאגיש טראכטן. **Time** ווייזט מיט 4 טשארטס פארוואס דער פארשריט וועט זיך נישט אפשטעלן.
- **8 אויגוסט:** **פאליטישע פארדריינגען, Hinton** – א פארשונג ווייזט אז **GPT-4** איז דער מערסט "לינקס-ליבערטארישער" מאדעל, און **LLaMA** דער מערסט "רעכטס-אויטאריטערער". **Geoffrey Hinton** מיינט אז AI וועט האבן געפילן.
- **14 אויגוסט:** **Andrew Ng וועגן וועלט-מאדעלן** – **Andrew Ng** זאגט אז מאדעלן ווי **Othello-GPT** בויען זיך אויף א בילד פון דער וועלט: ער האט אליין דערקענט דעם 8x8 ברעט און די כללים פונעם שפיל. דאס שטעלט אין פראגע **John Searle**'ס "Chinese Room" טענה.
- **15 אויגוסט:** **Gary Marcus: א פייל?** – **Gary Marcus** פרעגט צי גענעראטיוו AI וועט ארויסקומען א פייל. **Matt Clancy** און **Tamay Besiroglu** פירן א וויכוח וואס וועט געשען מיט דער עקאנאמיע ווען אינטעליגענץ וועט זיין גאר ביליג.
- **16 אויגוסט:** **זכרון-וואנט און Copilot לעקציעס** – אן ארטיקל ערקלערט אז **LLaMA.cpp** לויפט אויף א **MacBook** נישט פיל לאנגזאמער ווי אויף אן **A100**, ווייל ביים גענערירן איין טאקען נאך דעם אנדערן איז דער האלדער די שנעלקייט פונעם זכרון, און דער אונטערשייד דערין איז נישט אזוי גרויס. **GitHub** טיילט 10 כללים פון **Copilot** ווי צו בויען AI געצייג פאר פראגראמירער. אן אנדערער ארטיקל ווארנט אז

- אינהאלט געמאכט מיט AI וועט ווערן איינטעניג, און אז אמת'ע מענטשליכע ארבעט וועט זיין די רפואה. Tom Tunguz שרייבט פארוואס AI איז אויפן וועג צו איבערקערן ענינים.
- **17 אויגוסט: Ilya און OpenAI – generalization** הויפט-וויסנשאפטלער ערקלערט אין א לעקציע, לויט Kolmogorov complexity, אז נעוראלע נעצן זוכן פשטות. אויך: RLCF צו צופאסן קליינע מאדעלן דורך פראמפטינג, א טיפע לערנונג וועגן דער מאטעמאטיק פון טריינירן LLMs, א ביקורת "Does AI Just Suck?", און א שמועס אויף Hacker News צי מען קען זיך אליין אויסלערנען AI.
 - **18 אויגוסט: a16z און Lightspeed** a16z – אנאליזירט ווער וועט פארדינען דאס מערסטע פון AI אויף דער וואלקן- און אפליקאציע-שיכט. Lightspeed געבט פינף לעקציעס פאר גרינדער אין fintech.
 - **21 אויגוסט: "קיינ fine-tuning נישט נויטיג"** – אן ארטיקל טענהט אז מען דארף געווענליך נישט צופאסן א מאדעל מיט טריינירן, און אז few-shot prompting אדער RAG איז בעסער. עס איז אויך דא א ליסטע פון לייענען "קעגן דעם הייפ".
 - **22 אויגוסט: Elad Gil, SDXL, IBM** – Elad Gil זאגט אז די היינטיגע תקופה פון AI איז גאנץ אנדערש ווי פריער. Max Woolf האט פארבעסערט SDXL דורך אים טריינירן אויף שלעכטע בילדער וואס AI האט געמאכט. לויט IBM וועלן 40% ארבעטער דארפן אויסלערנען נייע פעאיגקייטן אין די קומענדיגע דריי יאר. א בלאג באשרייבט ווי מען שאפט טיילן פון 2D שפילן מיט אן RNN.
 - **23 אויגוסט: פאראויסזאגן, "פארקויפט ארבעט"** – א באריכט שאצט אפ צוויי יאר פאראויסזאגן וועגן פארשריט אויף בענטשמארקס. Sarah Tavel טענהט אז AI סטארטאפס זאלן פארקויפן די ארבעט אליין אנשטאט סאפטוועיר.
 - **24 אויגוסט: קרעאטיוויטעט און מאגערער אינטערעס** – אין א פארמעסט ביי Wharton פון אויסטראכטן אידייען, זענען 35 פון די 40 בעסטע אידייען געקומען פון ChatGPT-4. Ars Technica באשרייבט ווי GPT איז געווארן א געצייג פאר יעדע זאך. אן אנדער ארטיקל ווייזט אויף צאלן וואס דייטן אן אז דער אינטערעס פאלט: Bing געווינט נישט קיין מארק-אנטייל פון Google, און מען ניצט ChatGPT ווייניגער.
 - **25 אויגוסט: AI איז נאך נישט גוט גענוג** – SK Ventures טענהט אז AI איז נאך נישט גוט גענוג פאר די אויטאמאציע וואס די עקאנאמיע דארף. Mark Riedl טענהט אז די היינטיגע מאדעלן וועלן נישט ברענגען AGI, און אז די פיזיק אליין וועט נישט לאזן קיין פלוצעמדיגע שפרינגען.
 - **28 אויגוסט: Marc Andreessen** – Marc Andreessen זאגט אז AI וועט ברענגען גרויסע עקאנאמישע נוצן, און אז דער שרעק פאר אן AI וואס נעמט איבער איז איבערגעטריבן, ווייל זי האט נישט די טריבן וואס די עוואלוציע גיט.
 - **29 אויגוסט: גראף-פראמפטינג און כאאס** – אן איבערבליק פון פראמפטינג-מעטאדן וואס שטעלן פאר דאס טראכטן ווי א גראף. אן ארטיקל טענהט אז מאדעלן ברענגען כאאס, ווייל מען קען נישט וויסן פון פריער וואס זיי וועלן ארויסגעבן.
 - **30 אויגוסט: Llama 2 אזוי גוט ווי GPT-4 אין סיכומים** – Anyscale האט געפונען אז Llama-2-70b איז כמעט אזוי פאקטיש-גענוי ווי GPT-4 ביים מאכן סיכומים, אסאך בעסער ווי gpt-3.5-turbo, און אסאך ביליגער. אויך: אפשצן שמועס-מאדעלן ווערט מער און מער א געפיל, און מען האלט אז דער Open Assistant מאדעל איז פון די בעסטע. AI אין חינוך דארף מענטשליכע אויפזיכט, און Drew DeVault ברענגט א פעסימיסטישע מיינונג.

- **31 אויגוסט:** נייע יסוד-מאדעלן – Sebastian Raschka גייט טיף אריין אין Llama 2, Code Llama, GPT-4, פיין-טונינג אלס סערוויס, און קאפירעכט.

ווייטערע ידיעות

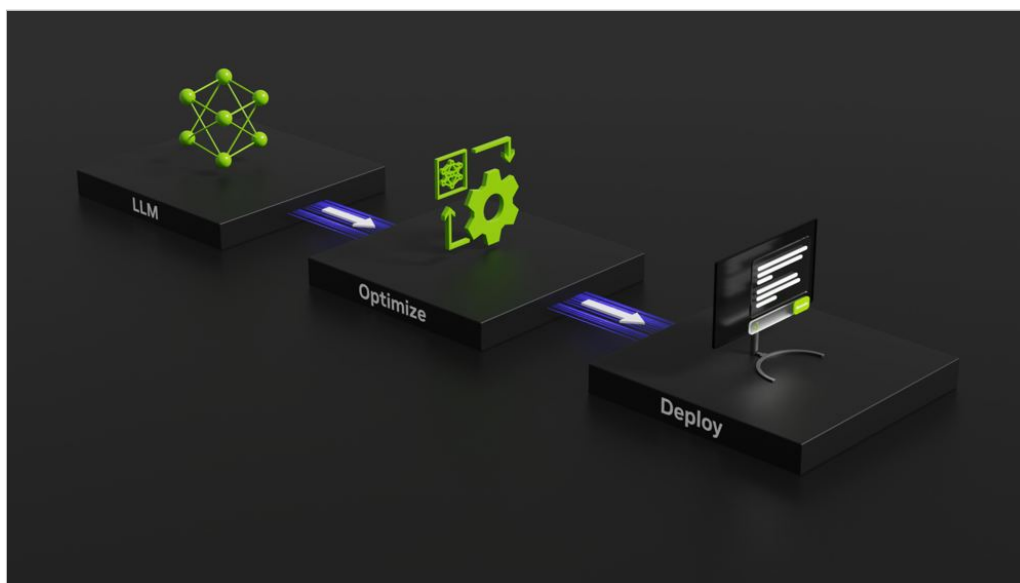
- **1 אויגוסט:** TechCrunch Disrupt – TechCrunch Disrupt (San Francisco, 21–19 סעפטעמבער, 2021) וועט אפגעבן א גאנצן טאג פאר AI.
- **10 אויגוסט:** Brain2Music – Google שאפט מוזיק פון בילדער פונעם מוח, און די מוזיק איז ענליך צו די לידער וואס דער מענטש האט געהערט ווען מען האט אים אונטערגעזוכט.
- **9 אויגוסט:** MeowLearning – א פירער פאר א סטארטאפ, פון דער פירערשאפט ביז די ML אינזשענירן, ווי צו ארבעטן מיט AI אין פראדוקציע.
- **1 אויגוסט:** עטליכע נייע אפליקאציעס – Musicfy טוישט א קול צו א געוואונטשענעם סטיל. Pump שפארט ביז 60% אויף BrainyPDF. AWS מאכט סיכומים פון PDF טעקעס. Fluent 2.0 איז א שפראך-לערער. Vizly אנאליזירט דאטא. SlideSpeak מאכט סיכומים פון PowerPoint פרעזענטאציעס.
- **2 אויגוסט:** ווייטערע געצייג – Unthread – קונה-סערוויס אין Slack; פייפליינס פאר Tagalog; Color-Diffusion – פארבט אן שווארץ-ווייסע בילדער; Artie – פירט דאטא אין רעאל-צייט פון דאטא-באזעס צו דאטא-לאגערן; Parea AI – געפינען און פאררעכטן פעלערן אין LLM אפליקאציעס; impel – א AI העלפער פאר Mac; Cedalio – א דאטא-באזע וואס מען קען נאכקוקן און נאכגיין פון וואו יעדע דאטא קומט.
- **14 אויגוסט:** נאך אפליקאציעס – Recast – סיכומים פון ארטיקלען אלס קול; Kyligence Copilot – אנאליז פון KPIs; ChartPixel – טשארטס; Consensus – זוכן אין פארשונג-ארבעטן; Interview Hero – סיכומים פון אינטערוויוס; Clay – זוכן קונים; Sum It Up – סיכומים; Kypso; Bloc; Dicer; Mine My Reviews; Gamma – פרעזענטאציעס אין ווייניגער ווי א מינוט; Algomo – קונה-סערוויס.

סעפטעמבער 2023 איז געווען א חודש פון גרויסע לאנסירונגען ביי ChatGPT, E 3-OpenAI: DALL, וואס קען זען, הערן און רעדן, און ווידער זוכן אויפ'ן אינטערנעט אין רעאלע צייט, צוזאמען מיט שמועסן וועגן א ווערט פון 80 ביז 90 ביליאן דאלאר. Amazon האט אנגעזאגט אז זי וועט אינוועסטירן ביז 4 ביליאן דאלאר אין Anthropic, און Microsoft האט פארגעשטעלט Microsoft Copilot פאר Windows 11 און Microsoft 365. אויף דער אפן-סאורס פראנט זענען ארויסגעקומען Falcon 180B, Persimmon-8B און דער ערשטער 7B מאדעל פון Mistral. אין וואשינגטאן האבן די גרויסע טעכנאלאגיע-פירער געשמועסט מיט סענאטארן וועגן רעגולירונג, נאך פירמעס האבן זיך אנגעשלאסן אינעם הסכם פון ווייסן הויז, און עס זענען צוגעקומען נאך קלאגעס פון שרייבער קעגן OpenAI.

מאדעלן און נייע לאנסירונגען

- **1 סעפטעמבער: Meta's DINO v2** – Meta האט ארויסגעלאזט FACET, א בענטשמארק וואס מעסט ווי "יושרדיק" עס זענען מאדעלן וואס דערקענען זאכן און מענטשן אין בילדער און ווידעאס. צוזאמען דערמיט האט זי איבערגעגעבן דעם DINO v2 מאדעל אונטער א ברייטערן לייסענס, אזוי אז די קהילה זאל קענען בויען אויף אים.
- **1 סעפטעמבער: Baidu's Ernie** – דער כינעזישער ריז Baidu האט אפיציעל ארויסגעלאזט זיין Ernie טשעטבאט, נאכדעם וואס די רעגירונג האט עס דערלויבט.
- **1 סעפטעמבער: Google TPU v5** – Google האט ארויסגעלאזט TPU v5 און מאשינען מיט H100 טשיפס. די טשיפס זענען שנעלער און די ווייכווארג-שטיצע איז בעסער, למשל פאר PyTorch, וואס איז אמאל געווען א שוואכקייט. לויטן נוולעטער מאכט דאס Google 'ס אנבאט צו א ממשות'דיקע קאנקורענץ פאר Nvidia.
- **5 סעפטעמבער: Ideogram** מיט **טיפאגראפיע** – דער סטארטאפ Ideogram, געשטיצט פון געוועזענע Google Brain פארשער, איז ארויס מיט 16.5 מיליאן דאלאר זוימען-געלט פון a16z און Index Ventures. זיין בילד-גענעראטאר שטעלט זיך אויס דערמיט וואס ער שרייבט פארלעסלעך טעקסט אינעם בילד.
- **6 סעפטעמבער: Hugging Face** טריינירונג-סערוויס – Hugging Face האט לאנסירט א סערוויס וואס טריינירט א מאדעל פאר'ן באניצער, אן דעם וואס ער זאל דארפן אליין פירן דעם קאמפליצירטן פראצעס.
- **6 סעפטעמבער: קאוד-מאדעל פון 1.6B** – Refact האט ארויסגעלאזט א מאדעל פון 1.6 ביליאן פאראמעטערס פאר שרייבן קאוד, וואס דערגרייכט 32% אויף HumanEval. ער איז טריינירט געווארן פון אנהייב אויף 1.2 טריליאן טאקענס, האלב קאוד און האלב שפראך.
- **7 סעפטעמבער: Falcon 180B** – דער נייער Falcon מאדעל פון די אראבישע עמיראטן האט 180 ביליאן פאראמעטערס און א קאנטעקסט פון 2k. ער איז א ביסל בעסער ווי Llama 2 70B. צוליב די גרויסע פארלאנגען פון ריסורסן איז אבער נישט קלאר צי די קהילה וועט אים אננעמען.

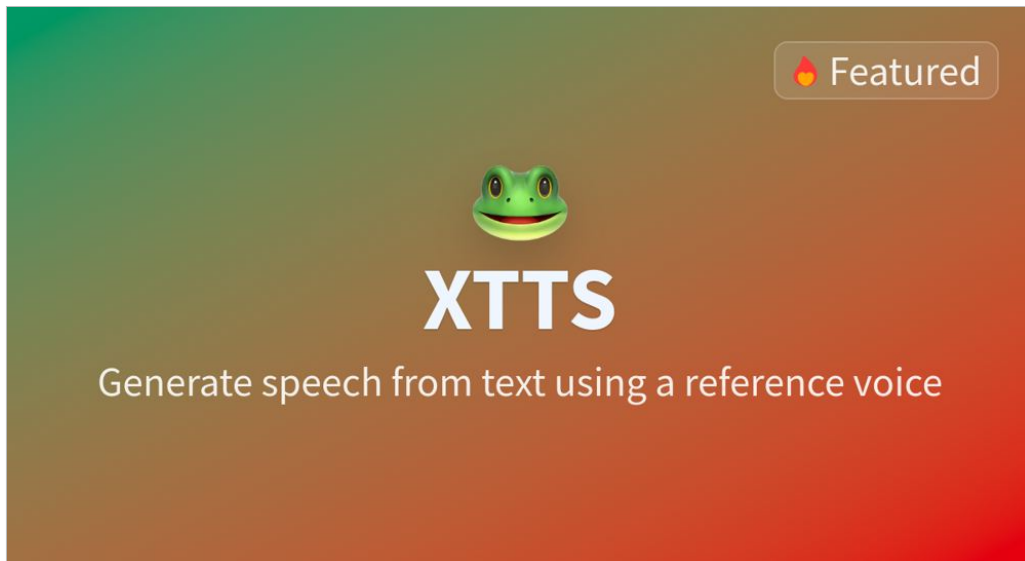
- **8 ספטמבער: Persimmon-8B פון Adept** – האט ארויסגעלאזט אן אפן-סאורס מאדעל פון 8 ביליאן פאראמעטערס מיט זייער גוטע רעזולטאטן, צוזאמען מיט קאוד פאר'ן לויפן אים. דער מאדעל האט 70 טויזנט אומבאנוצטע embeddings, וואס עפענען א טיר צו מולטימאדאלע ערווייטערונגען.
- **8 ספטמבער: Claude באצאלטער פלאן** – Anthropic האט לאנסירט א באצאלטע ווערסיע פון Claude פאר 20 דאלאר א חודש.
- **8 ספטמבער: IBM Watsonx** – IBM האט צוגעלייגט נייע גענעראטיווע מאדעלן און פעאיקייטן צו איר Watsonx פלאטפארמע.
- **11 ספטמבער: Nvidia TensorRT פאר שפראך-מאדעלן** – Nvidia האט ארויסגעלאזט א ווערסיע פון TensorRT וואס איז ספעציעל געמאכט פאר שפראך-מאדעלן אויף H100. דאס איז וויכטיג ווייל די קאסטן פון לויפן די מאדעלן וואקסן שנעל ווען מען שטעלט זיי ארויס פאר באניצער.



Nvidia TensorRT פאר שפראך-מאדעלן · 11 ספטמבער 2023

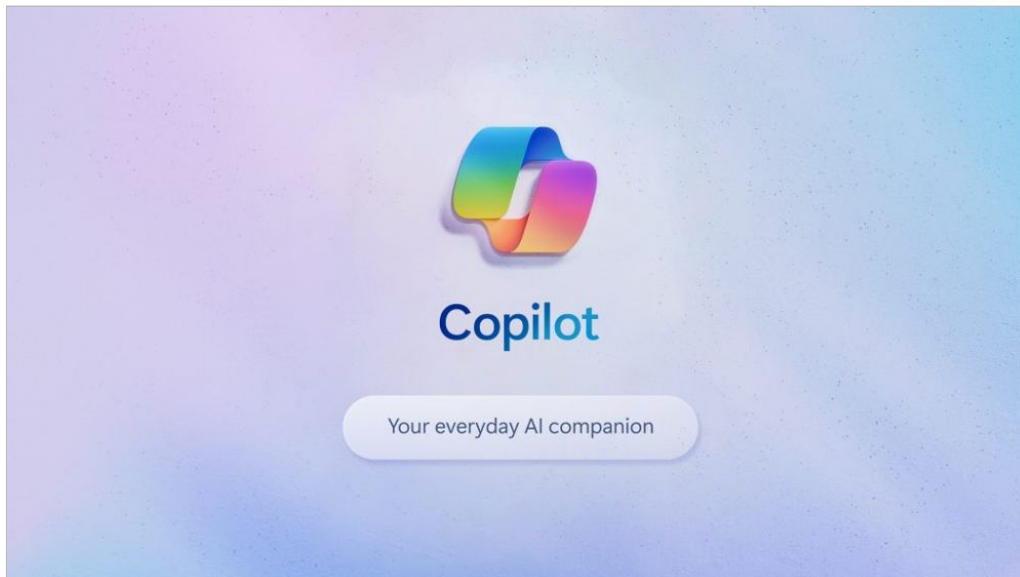
- **12 ספטמבער: Gutenberg אודיא-ביכער; Meta קעגן GPT-4** – Project Gutenberg האט צוזאמען מיט Microsoft ארויסגעגעבן טויזנטער אפענע ביכער אלס אודיא-ביכער, געמאכט מיט פארגעשריטענעם טעקסט-צו-רייד. באזונדער ווערט באריכטעט אז Meta קויפט אויף טשיפס און בויט דאטא צענטערס פאר א טשעטבאט וואס זאל קענען מיטהאלטן מיט GPT-4.
- **13 ספטמבער: OCR פאר אקאדעמישע דאקומענטן** – Facebook Research האט געוויזן א מאדעל וואס ציט ארויס טעקסט פון בילדער פון אקאדעמישע דאקומענטן, אריינגערעכנט מאטעמאטיק, מיט שטארקע רעזולטאטן. דאס קען העלפן דיגיטאליזירן אסאך אלטע טעקסטן.
- **14 ספטמבער: Stable Audio, Data Commons, Einstein Copilot** – Stability AI האט ארויסגעלאזט Stable Audio, וואס מאכט קורצע אודיא-שטיקלעך פון טעקסט-באפעלן. Google האט געגעבן Data Commons אן AI אינטערפעיס וואס נעמט אן פראגעס אין נאטירלעכער שפראך. Salesforce האט פארגעשטעלט Einstein Copilot מיט א "צוטרוי-שיכט" פאר זיכערהייט, פארבונדן מיט איר Data Cloud.

- **15 ספטמבער: EvoDiff און XTTS** – Microsoft האט געמאכט אפן-סאורס EvoDiff, א מאדעל וואס שאפט פארשידענע פראטיינען. Coqui האט ארויסגעגעבן די וואגן פון איר xtts מאדעל, וואס קען נאכמאכן א קול און רעדן אין פארשידענע שפראכן.



EvoDiff און XTTS · 15 ספטעמבער 2023

- **18 ספטעמבער: IBM MoE מאדעל; Adobe Firefly** – IBM האט ארויסגעלאזט עטליכע Mixture-of-Experts מאדעלן אלס אפן-סאורס, צוזאמען מיט זייער ארכיטעקטור און קאוד. Adobe's Firefly איז ארויס פון בעטא און שטייט צו דער רשות אין Creative Cloud, Express און Experience Cloud. מען באצאלט מיט "גענעראטיווע קרעדיטן", און ווען זיי ענדיגן זיך ארבעט עס ווייטער, נאר פאמעלעכער.
- **20 ספטעמבער: Bard עקסטענשאנס, ChatGPT, Anyscale – Bard** קען יעצט זוכן אינפארמאציע אין Gmail, Docs און Anyscale Drive. האט לאנסירט Endpoints פאר אפן-סאורס מאדעלן, און א פארבינדונג מיט Nvidia פאר ChatGPT. Ray'ס וויסן-גרעניץ איז אריבער פון ספטעמבער 2021 צו יאנואר 2022.
- **21 ספטעמבער: E 3-DALL און Copilot Chat** – OpenAI האט לאנסירט E 3-DALL, וואס פארשטייט פיל מער פיינהייטן און פרטים ווי פריערדיגע סיסטעמען, און וועט באלד זיין אריינגעבויט אין ChatGPT. GitHub האט געעפנט Copilot Chat פאר יחידים אין Visual Studio און Bard. VS Code האט באקומען א "דאבל-טשעק" פונקציע וואס פארגלייכט אירע ענטפערס מיט'ן אינטערנעט.
- **22 ספטעמבער: Microsoft Copilot** – Microsoft האט פארגעשטעלט Microsoft Copilot, אן AI געהילף אין Windows 11, Microsoft 365, Edge, וואס העלפט לויטן קאנטעקסט פון דעם וואס דער באניצער טוט. ער הייבט אן צו ווערן אויסגעטיילט אין Windows 11 דעם 26טן ספטעמבער. דער דערהיינטיגונג ברענגט אויך איבער 150 נייע AI פונקציעס אין אפליקאציעס ווי Paint און Photos.



2023 · Microsoft Copilot סעפטעמבער

- **25 סעפטעמבער: Bing מיט E 3-DALL** – Bing באקומט E 3-DALL, מער פערזענליכע ענטפערס, און געצייג וואס צייכנט בילדער אז זיי זענען געמאכט פון AI.
- **26 סעפטעמבער: ChatGPT זעט, הערט און רעדט** – מען וועט באלד קענען רעדן צו ChatGPT און הערן זיין ענטפער, און אויך ארויפלייגן בילדער און פרעגן וועגן זיי. Spotify האט ארויסגעלאזט אן איבערזעצונג פאר פאדקאסטס וואס באנוצט דאס אייגענע קול פונעם רעדנער, מיט Whisper פון OpenAI. Shortwave האט לאנסירט אן אימעיל-געהילף וואס זוכט אינעם גאנצן אימעיל-היסטאריע.
- **27 סעפטעמבער: Getty און YouTube** – Getty האט לאנסירט א בילד-גענעראטאר אויף Nvidia מיט שטארקע קאמערציעלע גאראנטיעס. YouTube האט פארגעשטעלט AI געצייג פאר שאפער, אריינגערעכנט Dream Screen פאר AI הינטערגרונטן אין Shorts, פארשלאגן פאר טעמעס, און מוזיק-רעקאמענדאציעס.
- **28 סעפטעמבער: ChatGPT; Mistral 7B זוכט ווידער** – Mistral האט ארויסגעלאזט איר ערשטן מאדעל, 7B, דורך א טארענט-לינק און מיט א פרייער לייסענס. די Instruct ווערסיע שטייט העכער ווי אסאך פון די בעסטע אינסטרוקציע-מאדעלן, "אן טריקס און אן פרייוואטע דאטא". ChatGPT קען ווידער זוכן אויפ'ן אינטערנעט דורך "Browse with Bing", צום ערשט פאר Plus און Enterprise. Meta לאנסירט אן אלגעמיינעם AI געהילף מיט Bing רעזולטאטן און בילד-שאפונג, און אויך פארשידענע AI כאראקטערן איבער אירע פלאטפארמעס.
- **29 סעפטעמבער: Meta Emu; Chroma מיט PaLM** – Meta Emu, Meta'ס בילד-מאדעל, איז פאר-טריינירט אויף 1.1 ביליאן בילד-פארן און דערנאך איינגעשטעלט אויף בלויז עטליכע טויזנט אויסגעקליבענע בילדער. ביי באניצער-אנקעטעס ווערט ער בעסער אפגעשאצט ווי SDXL, און ער שטייט אונטער סך פון Meta'ס נייעם געהילף. Chroma האט זיך פארבונדן מיט Google צו בויען אפליקאציעס אויף PaLM.

פארשונג און טעכניק

- **1 סעפטעמבער: בילד און ווידעא** – TexMod מישט אריין פרטים פון פארשידענע בילדער ביי שאפן בילדער פון ווייניג ביישפילן. MMVP באהאנדלט באזונדער באוועגונג און אויסזען ביים פאראויסזאגן ווידעא, און

איז גענויער און קאמפאקטער ווי די איצטיגע סיסטעמען. א דריטער פראיעקט פאסט צו דעם דעקאדער פאר אומבאקאנטע דאטאסעטן.

• **1 סעפטעמבער: פידבעק און חנפנות** – א פארשונג געפינט אז עס מאכט א גרויסן חילוק צי מען טריינירט מיט רעזינינגס אדער מיט ראנקינגס. אנדערע עקספערימענטן ווייזן אז OpenAI'ס באזע-מאדעלן לייזן נישט פון "חנפנות" (מסכים זיין מיט'ן מענטש אפילו ווען ער טעות'ט), כאטש Anthropic האט געוויזן אז גרעסערע מאדעלן זענען מער אזוי.

• **1 סעפטעמבער: דראון שלאגט האמפיאנען** – פארשער פון דעם אוניווערסיטעט פון ציריך און Intel האבן געבויט Swift, אן אויטאנאמע דראון וואס האט באזיגט מענטשלעכע וועלט-האמפיאנען אין FPV רעסינג.

• **4 סעפטעמבער: RLHF קעגן RLAI** – די ערשטע עפנטלעכע פארגלייכונג געפינט אז ביידע שיטות גיבן כמעט די זעלבע רעזולטאטן, און מענטשן ציען זיי פאר אין בערך 70% פון די פעלער. אנדערע ארבעטן: אן אלגעמיינע סיסטעם וואס נעמט אראפ עפעקטן ווי עכא פון אודיא, און א 3D שאפונג מיט מולטי-ווי דיפוזיע.

• **5 סעפטעמבער: Tiny Llama** – דער פראיעקט וויל טריינירן א Llama2 מאדעל פון 1.1 ביליאן פאראמעטערס אויף 3 טריליאן טאקעס. דאס איז פיל מער ווי עס לוינט זיך לויט די חשבונות פון טריינירן, אבער עס ברענגט א שטארקן פארטייל ביים לויפן דעם מאדעל. אויך ארויס: א סיסטעם צו לערנען נייע אויפגאבעס אן פארגעסן די אלטע, SpeechTokenizer פאר רייד, און א פארבינדונג פון קאמערע מיט IMU סענסארן צו כאפן מענטשלעכע באוועגונג.

• **6 סעפטעמבער: מאטעמאטיק ביי מאדעלן** – Rejection sampling Fine-Tuning (RFT) פארעסערט מאדעלן אין מאטעמאטיק, בפרט ווען מען נעמט ביישפילן פון עטליכע מאדעלן. אנדערע ארבעטן: SQLdepth פאר זעלבסט-פארנדיגע קארס, און א שיטה וואס שטעלט איין CLIP אויף א ספעציפישע אויפגאבע אן אנווערן זיין אלגעמיינע פעאיקייט.

• **7 סעפטעמבער: LoRA פאר PPO** – אנשטאט האלטן דריי פולע מאדעלן ביי RLHF, קען מען באנוצן LoRA אדאפטערס און שטארק פארמינערן דעם זכרון אן פארלירן אין רעזולטאטן. אויך ארויס: 3D פנים-אנימאציע געטריבן פון רייד, און אן "ארדער בוק" אין JAX וואס לויפט אויף GPU. fast.ai באריכטעט אז מאדעלן האבן אויסווענדעג געלערנט ביישפילן נאך זען זיי בלויז איין מאל.

• **8 סעפטעמבער: HyperDiffusion** – מען קען אנווענדן דיפוזיע גלייך אויף די וואגן פון א MLP און אזוי שאפן נייע 3D סצענעס. אנדערע ארבעטן: אנאליז פון יערליכע באריכטן פון פירמעס כדי פאראויסצוזאגן אקציע-פרייזן, און איין ברייטע feedforward שיכט פאר'ן ענקאדער און דעקאדער צוזאמען, וואס פארשנעלערט שטארק דאס לויפן מיט ווייניג פארלוסט.

• **11 סעפטעמבער: 3D און רעקאמענדאציעס** – א פארשונג געפינט ווו א באסקעטבאל איז אין 3D פון איין בילד. אן אנדערע פרובירט ווי גוט ChatGPT רעקאמענדירט פילמען און ביכער. א דריטע שאפט 3D אוואטארן מיט קליידער פון 2D normal maps.

• **12 סעפטעמבער: Medusa** – אנשטאט א באזונדערן קליינעם "דראפט מאדעל" פאר speculative decoding, דינט Medusa אלס דראפט פאר זיך אליין. אויך ארויס: NuPrompt, באפעלן אין פשוטער שפראך פאר זעלבסט-פארנדיגע קארס, און א בענטשמארק פאר לאגישע טראכטן ביי מאדעלן פון זעאונג און שפראך.

- **13 סעפטעמבער: AI שמעקט** – פארשער האבן טריינירט א גראף נעטווארק אויף טויזנטער מאלעקולן וואס זענען געצייכנט געווארן מיט דער האנט, און עס דערקענט ריחות גענוי. אנדערע ארבעטן: צי AI שרייב-געהילפן ווי InstructGPT מאכן דאס שרייבן פון מענטשן מער ענליך איינס צום אנדערן, און Scenimefy וואס מאכט פון פאטאס אנימע-בילדער.
- **14 סעפטעמבער: UMMAFormer און AnomalyGPT** – UMMAFormer געפינט די געפעלטע טיילן אין א ווידעא. AnomalyGPT געפינט פעלער אין פאבריקן אן מאנועלע איינשטעלונגען. א דריטע פארשונג געפינט אז מאדעלן געדענקען יסוד-וויסן נישט שלעכט, אבער זיי זענען נישט גאנץ ביים טראכטן דערמיט.
- **15 סעפטעמבער: אליינמענט אן טריינירן** – ווען מען לאזט א מאדעל זיך צוריקדרייען ביים שרייבן נאכן אפשאצן זיין אייגענעם ארויסגאנג, פארבעסערט זיך די אליינמענט פון א "פארפרוירענעם" מאדעל ביז 81%. אויך ארויס: ביליגערע איינשטעלונג פון MoE מאדעלן, און HAMUR פאר רעקאמענדאציעס איבער פארשידענע געביטן.
- **18 סעפטעמבער: קאפירעכט איבערבליק** – א פארשונג-ארבעט פאסט צונויף די קאפירעכט-פראגעס ביי AI, איינגעטיילט אין פראגעס פון אריינגאנג און פון ארויסגאנג, אן צו געבן א מיינונג. אנדערע ארבעטן: EncodecMAE, איין אודיא-מאדעל פאר רייד, מוזיק און קלאנגען, און 40 SyntheWorld, טויזנט סינטעטישע בילדער פון דער ערד פון אויבן.
- **19 סעפטעמבער: SCT און ווידעראשפּרוּכּן** – Salient Channel Tuning קלייבט אויס וועלכע טיילן פון א מאדעל צו איינשטעלן, מיט פיל ווייניגער פאראמעטערס. מאדעלן ווערן צעמישט ווען זיי טרעפן אינפארמאציע וואס שטימט נישט מיט וואס זיי ווייסן שוין. א נייער ציל פאר טריינירן, וואס בינדט צו embeddings מיט א contrastive loss, איז שנעלער און גענויער.
- **20 סעפטעמבער: CulturaX** – א גערייניגטער דאטאסעט פון בערך 6 טריליאן טאקענס אין פילע שפראכן. אויך ארויס: א שיטה צו צעטיילן לאנגע טעקסטן פאר שאפן בילדער, און טיפקייט-שאצונג וואס טיילט אפ די קאמערע-פרטים פונעם אינהאלט.
- **21 סעפטעמבער: רובאטן לערנען איבער נאכט** – דער Toyota Research Institute לערנט א רובאט א נייע פעאיקייט איבער איין נאכט מיט דיפּוּזע-מאדעלן. זיין סיסטעם איז שוין טריינירט אויף 60 פעאיקייטן. אנדערע ארבעטן: RenderIH און TransHand פאר הענט-פאזיציעס, רעקאמענדאציעס מיט שפראך-מאדעלן, און מעדיצינישע ענטפערס.
- **21 סעפטעמבער: OPRO** – Google's OPRO פארבעסערט מאטעמאטיק ביי מאדעלן דורך מענטשלעכע ערמוטיגונג אינעם באפעל, למשל "נעם א טיפן אטעם" אדער "לאמיר טראכטן שריט ביי שריט".
- **22 סעפטעמבער: UNet פארבעסערונג** – מען קען פארבעסערן Stable Diffusion אן נאך טריינירן, דורך באוועגן וואגן און איבערגרייסן residuals אינעם UNet. אויך ארויס: AV-SUPERB פאר קלאנג און בילד צוזאמען, און WAKD פאר דיסטילאציע.
- **25 סעפטעמבער: "A איז B"** – Gary Marcus שרייבט אז מאדעלן וואס זענען טריינירט אויף "A איז B" קענען נישט פון זיך אליין אפלערנען אז "B איז A". אנדערע ארבעטן: EmoPP פאר רייד מיט געפילן, א מאדעל פון מענטשלעכע קערפערס טריינירט בלויז אויף סינטעטישע דאטא, און TCOVIS פאר ווידעא.
- **26 סעפטעמבער: 3D און דאקומענטן** – NeuralLabeling און PVLFF צייכענען 3D סצענעס מיט טעקסט. LayoutLM איז געווארן בעסער אין ליינענען ביזנעס-דאקומענטן מיט צוויי נייע טריינירונג-אויפגאבעס.

- **27 סעפטעמבער: NeRF און ווידעא** – א נייער "refractive-reflective field" באהאנדלט שיינענדיגע און דורכזיכטיגע זאכן. אויך ארויס: קאנסיסטענטער בילד-באארבעטונג, און BSCV פאר פאריכטן צעבראכענע ווידעאס.
- **28 סעפטעמבער: RankVicuna, QA-LoRA, Llava-RLHF** – RankVicuna איז אן אפן-סאורס געצייג וואס סארטירט זוך-רעזולטאטן מיט שפראך-מאדעלן. QA-LoRA מאכט מאדעלן קלענער און שנעלער. Llava איז דורכגעגאנגען אינסטרוקציע-טריינירונג און RLHF, און פאלגט בעסער אינסטרוקציעס.
- **29 סעפטעמבער: VQ-VAE און JAM** – א פשוטע קוואנטיזאציע-שיטה פארמיידט codebook collapse און דארף נישט די קאמפליצירטע מעכאניזמען. JAM פארבינדט באזונדערע מאדעלן, און דערגרייכט שטארקע רעזולטאטן אין מולטימאדאלע אויפגאבעס.

געצייג פאר פראגראמירער און אפן-סאורס

- **4 סעפטעמבער: געצייג** – Gaussian painting, א סטערעאגראפישער עפעקט פון קונסטווערק. Modular Diffusion, א מאדולארער API פאר דיפוזיע אין PyTorch. TextBase, א פשוטער ראם פאר טשעטבאטס. אויך אן ארטיקל ווי אזוי צו בויען פלאגיןס מיט WASM.
- **5 סעפטעמבער: פיינ-טיונינג און RAG** – LLM Fine-tuning Hub; ragas צו אפשצן RAG פייפליינס; דער LLM CLI געצייג פון Simon Willison מיט embeddings; און GPT Researcher, אן AI אגענט וואס פארשט אנליין.
- **6 סעפטעמבער: LoRA באטשינג** – מען קען לויפן פארשידענע LoRA ווערסיעס אויף איין GPU און דינען פארשידענע באניצער. אויך: Topical-Chat, א דאטאסעט פון שמועסן אויף 8 טעמעס, און SAM.cpp, א Meta's Segment Anything אין C++.
- **7 סעפטעמבער: TokenFlow** – TokenFlow באארבעט ווידעאס גלאטער, און לויטן נוזלעטער כמעט בעסער ווי Runway Gen-2. אויך: Comgra צו קוקן אינעווייניג אין נעטווארקס, ReliTalk פאר אוואטארן, און Doppelgangers פאר 3D פארגלייכונג.
- **8 סעפטעמבער: OnPrem.LLM** – א Python פעקל צו לויפן מאדעלן אויף סענסיטיווע דאטא און אן אינטערנעט. אויך: ResFields פאר 3D סצענעס וואס טוישן זיך מיט דער צייט, און מאדעלן פאר בילד-רעסטאוראציע.
- **11 סעפטעמבער: SAM-Med2D** – SAM טריינירט אויף מעדיצינישע בילדער ארבעט פיל בעסער אויף זיי. אויך: XenonJS אן קאוד, און Web AI צו לויפן מאדעלן אינעם בראווער.
- **12 סעפטעמבער: RestGPT און PointLLM** – RestGPT איז אן AI אגענט וואס קאנטראלירט אפליקאציעס. PointLLM לערנט מאדעלן צו פארשטיין 3D זאכן. Eugene Yan ברענגט זיבן פראקטישע מוסטערן פאר בויען סיסטעמען מיט שפראך-מאדעלן.
- **13 סעפטעמבער: Arcus און Whisper Turbo** – Arcus טיילט איין דאקומענטן אין גרופעס כדי צו זוכן איבער ריזיגע מאסן דאטא מיט ווייניגער האלוצינאציעס. Whisper Turbo, געשריבן אין Rust, ערעצט דעם Whisper API און איז 20 מאל שנעלער.

- **14 סעפטעמבער: ExLlamaV2** – א ביבליאטעק צו לויפן לאקאלע מאדעלן אויף געווענליכע GPUs . אויך: Rectified Flow צו פארשנעלערן טעקסט-צו-בילד, און א פארשונג וואס ווייזט אז לערנען איין פראגראמירונג-שפראך העלפט ביי אנדערע.
- **15 סעפטעמבער: RAG פירער** – א פולער פירער פאר RAG אפליקאציעס פאר פראדוקציע. אויך: DeepFakeFace דאטאסעט, און JPEG וואס פאסט פאר דיפ-לערנען.
- **18 סעפטעמבער: Repilot** – Repilot מאכט קאד-קאפילאטן בעסער אין פאריכטן באגס. אויך: א סטאבילער מולטי-ווי 3D מאדעל, טראנספארמערס פאר EEG שלאף-אנאליז, און Illusion diffusion פאר ספיראל-בילדער.
- **19 סעפטעמבער: VimGPT** – אן עקספערימענטאלער אגענט מיט דירעקטן צוטריט צו Vim. אויך: 2D מיט 3D צוזאמען, און Robust e-NeRF פאר event קאמערעס.
- **20 סעפטעמבער: DoLa** – DoLa פארגלייכט שיכטן אינעם מאדעל כדי ער זאל זיך האלטן ביי די פאקטן. אויך: PIPS2 וואס פאלגט נאך הונדערטער פונקטן אין רעאלע צייט, LayoutNUWA פאר דיזיין, און ViTMatte פאר אראפנעמען הינטערגרונט.
- **21 סעפטעמבער: OpenPipe** – OpenPipe מאכט פון טייערע באפעלן ביליגע מאדעלן. אויך: DFormer פאר RGB-D, און אן ארטיקל אז RAG דארף מער ווי בלויז embedding זוך.
- **22 סעפטעמבער: DreamLLM** – DreamLLM שאפט טעקסט און בילדער צוזאמען. אויך: א פירער פאר ML אינוועניריע, שנעלע אנאליז פון 3D פנים, און וויסן-גראפן מיט 50 שורות קאד.
- **25 סעפטעמבער: Flash Attention 2** – Flash Attention 2 ווערט אריינגעבויט אין HuggingFace. אויך: Sidekick פאר פראגראמירער, FGFusion פאר קאמערע און Lidar, שנעלערע embeddings, און א קורס פאר עפעקטיוון ML.
- **26 סעפטעמבער: Odin און DeepEval** – Odin מאכט וויסן-גראפן מיט DeepEval. GPT-4. שאצט אפ LLM פייפליינס. אויך: סימולירטע אקטיווע סטערעא פאר טיפקייט.
- **27 סעפטעמבער: LAION OpenLM** – א ביבליאטעק צו טריינירן מאדעלן ביז 7 ביליאן פאראמעטערס, מיט מאדעלן וואס זענען טריינירט אויף 1.25 טריליאן טאקענס. אויך: אנאטירטע פארשונג-ארבעטן, און קאטעגאריע-טעאריע פאר ML.
- **28 סעפטעמבער: DRLX** – א ביבליאטעק פאר RL אויף דיפוזיע-מאדעלן. אויך: HMTF-Net פאר בילד-פארבעסערונג, און UnitedHuman.
- **29 סעפטעמבער: SigLIP** – Google האט ארויסגעגעבן SigLIP checkpoints און דערהיינטיגטן קאד. אויך: InternLM-XComposer, און Metaphor ס'API וואס פארבינדט מאדעלן מיט'ן אינטערנעט.

פירמעס, געלט און ביזנעס

- **4 סעפטעמבער: Contextual און EvolutionaryScale** – Contextual האט אויפגעהויבן 20 מיליאן דאלאר פאר RAG ביי פירמעס. EvolutionaryScale, געגרינדעט פון געוועזענע Meta פארשער, האט אויפגעהויבן 40 מיליאן פון Lux Capital ביי א ווערט פון 200 מיליאן, פאר פאראויסזאגן פראטיין-סטרוקטורן. זי טענהט אז זי איז שנעלער ווי AlphaFold.

- **6 סעפטעמבער: Kindo – Kindo** האט אויפגעהויבן 7 מיליאן, אנגעפירט פון Riot Ventures, פאר געצייג צו פירן AI פראצעסן ביי פירמעס.
- **7 סעפטעמבער: Apple און DevDay** Apple – גיט לויט באריכטן אויס מיליאנען דאלאר יעדן טאג אויף AI פאר Siri. OpenAI וועט האלטן איר ערשטע DevDay קאנפערענץ אין סאן פראנציסקא דעם 6טן נאוועמבער. מער ווי 2 מיליאן פראגראמירער באנוצן שוין אירע געצייג.
- **8 סעפטעמבער: AI Grant ; Microsoft גאראנטיע** – AI Grant האט אנגעזאגט איר צווייטע גרופע סטארטאפס, אריינגערעכנט Microsoft GGML. זאגט אז זי וועט נעמען די לעגאלע פאראנטווארטליכקייט אויב Copilot באניצער ווערן געקלאגט פאר קאפירעכט.
- **11 סעפטעמבער: Imbue – Imbue** האט אויפגעהויבן 200 מיליאן אין Series B ביי א ווערט פון איבער 1 ביליאן, פאר סיסטעמען וואס טראכטן און שרייבן קאוד. NERSC גיט 50% הנחה אויף איר A100 סופערקאמפיוטער דורכאויס סעפטעמבער. Microsoft און Paige בויען דעם גרעסטן בילד-מאדעל אויף דער וועלט צו דערקענען קענסער.
- **13 סעפטעמבער: Google פאנד; MLCommons** – Google.org גיט 20 מיליאן פאר Digital Futures Project פאר פאראנטווארטליכע AI. אויף אן MLCommons בענטשמארק פאר שנעלקייט האט Nvidia פארנומען דאס ערשטע ארט, און Intel דאס צווייטע.
- **14 סעפטעמבער: Moderna, Amazon** – דער הויפט פון Moderna דערציילט אז די פירמע באנוצט AI שוין פון פאר ChatGPT, אויך ביים שאפן די COVID-19 וואקסין. Amazon מוטיגט פארקויפער צו מאכן פראדוקט-באשרייבונגען מיט AI. Lemonade's "Seeds" האט אין ווייניגער ווי א חודש געהאלפן אנהייבן איבער 7,000 לידער.
- **15 סעפטעמבער: Patronus און MLPerf** Patronus AI האט אויפגעהויבן 3 מיליאן, אנגעפירט פון Lightspeed, צו אפשאצן מאדעלן פאר פירמעס. MLPerf Inference v3.1 האט צוגעלייגט א שפראך-מאדעל צו די טעסטן, און MLPerf Storage v0.5 מעסט סטארעדזש ביים טריינירן.
- **19 סעפטעמבער: Writer און Databricks** – Writer האט אויפגעהויבן 100 מיליאן אין Series B, אנגעפירט פון ICONIQ Growth, און זי טריינירט נישט אויף די דאטא פון אירע קונים. Databricks האט אויפגעהויבן איבער 500 מיליאן ביי א ווערט פון 43 ביליאן, מיט Nvidia צווישן די אינוועסטארן. Google בויט מיט'ן פענטאגאן אן AI מיקראסקאפ קעגן קענסער, און GE בויט אן AI אולטראסאונד.
- **22 סעפטעמבער: Kevin Systrom – Systrom**, איצט ביי Artifact, גלייבט אז AI וועט מאכן מענטשן "סופער-שטארק" און נישט שאטן.
- **25 סעפטעמבער: Secoda – Secoda** האט אויפגעהויבן 14 מיליאן, אנגעפירט פון Craft Ventures. אירע איינקונפטן וואקסן 250% א יאר.
- **26 סעפטעמבער: Amazon אין Anthropic** – Amazon אינוועסטירט ביז 4 ביליאן דאלאר אין Anthropic, מיט אפמאכן איבער IP און קאמפיוטער-קראפט. Anthropic'ס זיכערהייט-אויפזיכט בלייבט אומגעטוישט. Forethought האט לאנסירט Autoflows.



2023 Anthropic אין Amazon · 26 סעפטעמבער

- **27 סעפטעמבער: OpenAI ווערט** – OpenAI שמועסט וועגן פארקויפן אקציעס, וואס וואלט געהויבן איר ווערט פון 29 ביליאן צו 80 ביז 90 ביליאן. די CIA בויט א טשעטבאט אויף עפנטלעכע דאטא.
- **28 סעפטעמבער: Jony Ive מיט Sam Altman** – לויט באריכטן קוקן דער געוועזענער הויפט-דיזיינער פון Apple און Altman אויף צו בויען צוזאמען אן AI מכשיר.
- **29 סעפטעמבער: Autoflows, Anthropic מיט BCG** – Autoflows, Anthropic, SupportGPT, באשליסט אליין די שריטן פאר קונה-סערוויס. Anthropic פארבינדט זיך מיט BCG צו אנבאטן Claude פאר BCG'ס קונים.

רעגירונגען, געזעצן און זיכערהייט

- **4 סעפטעמבער: X און דאטא** – X'ס נייע פריוואטקייט-פאליסי זאמלט ביאמעטרישע דאטא, אינפארמאציע וועגן ארבעט און חינוך, און וועט באנוצן עפנטלעכע דאטא צו טריינירן מאדעלן. Musk זאגט אז נאר עפנטלעכע דאטא וועט באנוצט ווערן.
- **11 סעפטעמבער: Meta אפציע** – Meta האט אריינגעשטעלט א פארם וואס לאזט ביטן אראפנעמען פערזענליכע דאטא פון דריטע זייטן פון טריינירן.
- **12 סעפטעמבער: קלאגעס קעגן OpenAI** – נאך א גרופע שרייבער קלאגט OpenAI פאר טריינירן אויף זייערע ווערק. OpenAI ענטפערט אז דאס איז א "טראנספארמירנדיגע באנוץ", ברענגט אלס משל דעם Google Books פאל, און בעט אפּווארפן רוב טענות. OpenAI באשטעטיגט אויך אז AI שרייב-דעטעקטארן ארבעטן נישט.
- **13 סעפטעמבער: ווייסן הויז הסכם** – Adobe, IBM, Nvidia און אנדערע האבן זיך אנגעשלאסן אינעם פרייוויליגן הסכם פון ווייסן הויז, וואס באטאנט טעסטן פאר'ן ארויסלאזן און טיילן ריזיקעס.
- **18 סעפטעמבער: טעכנאלאגיע-פירער ביי סענאטארן** – Musk, Gates, Zuckerberg און די גרינדער פון Anthropic און OpenAI האבן געשמועסט מיט סענאטארן אין וואשינגטאן. Musk האט געווארנט פון סכנה, Gates האט גערעדט וועגן קעמפן קעגן הונגער, און עס איז ארויסגעקומען א הסכמה אז די רעגירונג דארף האבן אויפזיכט.

- **20 סעפטעמבער: OpenAI – Red Teaming** גרינדעט א נעץ פון עקספערטן צו אפשאצן און פארמינערן ריזיקעס פון אירע מאדעלן.
- **29 סעפטעמבער: קלאוד מעלדונגען** – דער ווייסן הויז קען פארלאנגען אז קלאוד פירמעס זאלן מעלדן גרויסע קויפן פון קאמפיוטער-קראפט, כדי צו כאפן אויסלענדישע סכנות. Microsoft און OpenAI שטיצן ענליכע אויפזיכט. Google-Extended'ס Google לאזט וועבזייטלעך זיך ארויסנעמען פון טריינירן, אבער ווייטער ערשיינען אין Search.

אנאליזן און מיינונגען

- **1 סעפטעמבער: פובליקום און פארלאנגזאמונג** – א שרייבער טענהט אז די אנקעטעס וואס ווייזן אז רוב מענטשן ווילן פארלאנגזאמען AI האבן פראבלעמען.
- **4 סעפטעמבער: ChatGPT באנוץ** – SparkToro געפינט אז דער פארקער ביי OpenAI איז געפאלן 29% זינט מאי. בערך 29% פון די באפעלן זענען וועגן פראגראמירן און 10% וועגן חינוך.
- **5 סעפטעמבער: אליינמענט** – אן אינטערוויו מיט'ן הויפט פון אליינמענט-פארשונג ביי OpenAI.
- **6 סעפטעמבער: עטישע פראגעס** – אן ארטיקל וועגן די עטישע דילעמעס פון AI.
- **7 סעפטעמבער: מאכט-איבערטיילונג** – אן ארטיקל אין Time טענהט אז AI וועט ברענגען אן היסטארישע איבערטיילונג פון מאכט.
- **8 סעפטעמבער: Musk היסטאריע** – א צוריקבליק אויף Musk'ס אקטיוויטעטן אין AI אין די פריע 2010ער יארן.
- **11 סעפטעמבער: Greylock און OpenAI** – Greylock זאגט אז אינוועסטארן ווענדן זיך פון Nvidia צו פראדוקטן וואס זענען פון יסוד אויף AI. Wired שרייבט אז די יאגעניש נאך געלט און פראדוקטן האט געטוישט OpenAI'ס קולטור. קריטיקער טענהן אז דאס נעמט אפ די אויפמערקזאמקייט פון ריזיקעס, און OpenAI זאגט אז איר מיסיע איז די זעלבע.
- **12 סעפטעמבער: קענען מאדעלן טראכטן?** – די ויכוח צי מאדעלן טראכטן אדער נאר געדענקען, און די רעזולטאטן פון די פארשונגען זענען געמישט.
- **13 סעפטעמבער: פערזענליכע סערוויסעס; רעגולירונג** – a16z זעט א ריזיגן מארק פאר פערזענליכע אנליין סערוויסעס. אן אנדער ארטיקל ערקלערט פארוואס עס איז שווער צו רעגולירן AI.
- **14 סעפטעמבער: SaaS** – אן אנאליז פון די השפעה פון גענעראטיווע AI אויף עקזיסטירנדיגע SaaS פירמעס.
- **15 סעפטעמבער: גלאבאלע אויפזיכט** – די קאנקורענץ צווישן אמעריקע און כינע מאכט שווער א וועלטלעכן הסכם איבער AI.
- **18 סעפטעמבער: קאנסולטאנטן מיט GPT-4** – א פארשונג געפינט אז קאנסולטאנטן וואס באנוצן GPT-4 שניידן אסאך בעסער אפ ווי די וואס נישט. Lightspeed זעט זיכערהייט פון AI אלס א גרויסע געלעגנהייט. אן אנדער ארטיקל טענהט אז GPT-4 ווערט נישט ערגער.

- **19 סעפטעמבער: אפן-סאורס וועט געווינען** – א שרייבער טענהט אז קאנטראל און צוטרוי זענען וויכטיגער ווי שכל, און דערפאר וועט אפן-סאורס געווינען. Elad Gil טענהט אז עס איז נאך צו פרי צו רעגולירן. אן אנדער ארטיקל טענהט אז AI וועט נישט מאכן מענטשלעכע שעפערשישקייט איבעריג.
- **21 סעפטעמבער: Sequoia "צווייטער אקט"** – Sequoia זאגט אז גענעראטיווע AI גייט אריבער צו ליין מענטשלעכע פראבלעמען פון אנהייב ביזן סוף. VersionOne פרעגט צי מען האט שוין געפונען דעם ביזנעס-מאדעל וואס פאסט פאר AI.
- **22 סעפטעמבער: Altman און ארבעט** – אן אינטערוויו מיט Sam Altman איבער רעגולירונג און די צוקונפט פון OpenAI. אן ארטיקל פרעגט צי AI קען ראטעווען די נייעס-אינדוסטריע. לויט Gallup גלייבן 75% אמעריקאנער אז AI וועט פארמינערן די צאל דזשאבס אין די קומענדיגע 10 יאר.
- **25 סעפטעמבער: Prompt engineering** – א שרייבער טענהט אז E 3-DALL שטעלט אין סכנה די צוקונפט פון "prompt engineering".
- **26 סעפטעמבער: אינטערפעיסעס אין 2027** – א שרייבער טענהט אז גענעראטיווע AI וועט נישט ערוועצן די פריערדיגע UX כידושים, נאר זיך צוגעבן צו זיי.
- **27 סעפטעמבער: GPT-4 מיט זעאונג; 200 ביליאן פראגע** – GPT-4 מיט זעאונג האט נאך שוואכקייטן און האלוצינירט. Sequoia שרייבט אז סטארטאפס דארפן זיך ווענדן פון אינפראסטרוקטור צו ריכטיגן ווערט פאר'ן קונה.
- **28 סעפטעמבער: פירמעס** – Big Technology שרייבט אז גענעראטיווע AI איז זייער נוצלעך פאר פירמעס, אבער עס וועט נעמען צייט.
- **29 סעפטעמבער: SaaS מיט a16z** – א שמועס מיט a16z, אז AI פאר SaaS איז מער וועגן טובנות ווי וועגן אינפארמאציע.

ווייטערע ידיעות

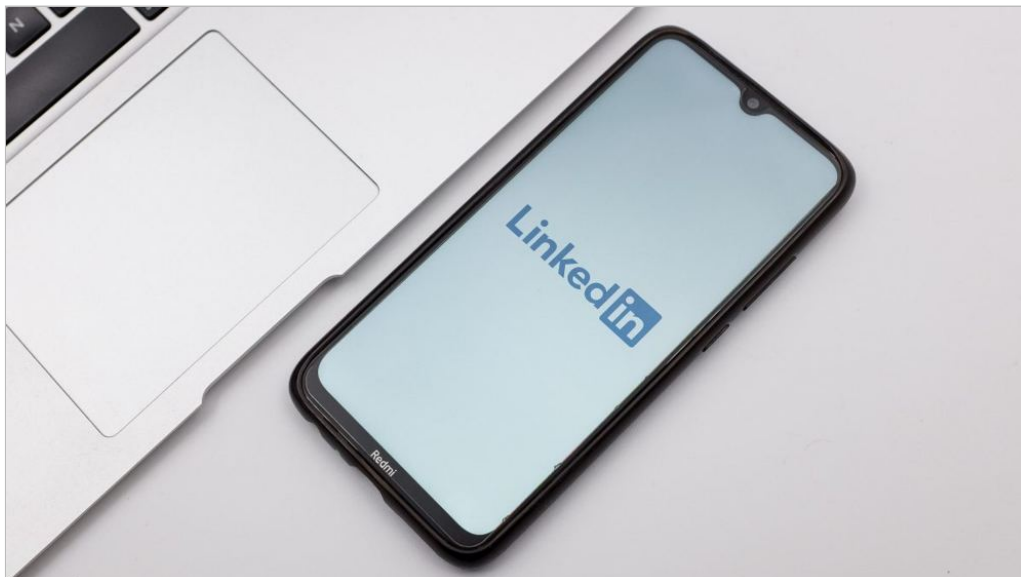
- **1 סעפטעמבער: ai פאר Anguilla** – דער קליינער אינדזל Anguilla קען פארדינען ביז 30 מיליאן דאלאר דעם יאר פון דעם "ai." דאמעין.
- **5 סעפטעמבער: OpenAI פאר לערער; Gizmodo** – OpenAI האט ארויסגעגעבן א פירער פאר לערער ווי אזוי צו באנוצן ChatGPT אין קלאס. G/O Media האט פארמאכט olñGizmodo en Espa און באנוצט אנשטאט דעם AI איבערזעצונגען, וואס האבן געמישט שפאניש מיט ענגליש. עס איז אויך דא אן ארטיקל וועגן AI און שמועסן מיט חיות.
- **6 סעפטעמבער: קאלעדזש עסייען** – קאלעדזשעס ראנגלען זיך מיט עסייען געשריבן מיט AI. Georgia Tech ערלויבט עס צו באנוצן פאר אידעייען, אבער נישט צו שרייבן דעם אינהאלט.
- **7 סעפטעמבער: Kaedim** – דער 3D סטארטאפ Kaedim האט אפט באנוצט מענטשלעכע קינסטלער, טייל מאל צו מאכן די מאדעלן אינגאנצן.
- **8 סעפטעמבער: ProteinQure** – ProteinQure האט מיט ML צוגעגרייט PQ-001, א מעדיצין קעגן ברוסט-קענסער וואס צילט נאר די קראנקע צעלן, נאך געלונגענע טעסטן אויף חיות. דער "Ghostwriter" האט ארויסגעגעבן נאך א ליד מיט AI קולות.

- **11 סעפטעמבער: וואסער - ChatGPT** באנוצט ביז 500 מיליטער וואסער פאר יעדע 5 ביז 50 באפעלן.
- **13 סעפטעמבער: Roblox - Roblox** האט פארגעשטעלט אן AI געהילף פאר שאפן וועלטן און 3D מאדעלן, וואס איז ערווארט שפעט דעם יאר אדער פרי נעקסטן יאר.
- **15 סעפטעמבער: RETFound** - א מאדעל וואס דיאגנאזירט פון בילדער פון די אויגן-רעטינע אויגן-קראנקהייטן, הארץ-פראבלעמען און ריזיקע פון Parkinson.
- **20 סעפטעמבער: AlphaFold - AlphaFold** האט געהאלפן באשטימען סטרוקטורן פון צוויי "אומאָרגאניזירטע" פראטיינען ווען זיי ווירקן איינס אויפ'ן אנדערן.
- **21 סעפטעמבער: Microsoft אנאנס - Microsoft** האט פלאנירט אן אונטערנעמונג צו פארשטעלן איר AI פלאן פאר Windows און 365.
- **26 סעפטעמבער: ChatGPT דיאגנאזירט** - נאך דריי יאר און 17 דאקטוירים האט ChatGPT פארגעשלאגן tethered cord syndrome ביי א יונגל. עקספערטן ווארענען אבער צו זיין פארזיכטיג.
- **27 סעפטעמבער: פאעטן - AI** פירמעס דינגען אויף פאעטן און שרייבער צו פארבעסערן די קוואליטעט פון דעם טעקסט וואס זייערע געצייג שרייבן.
- **28 סעפטעמבער: ספיראל קונסט - ספיראל-בילדער** געמאכט מיט Stable Diffusion און ControlNet זענען געווארן וויראל.
- **1 סעפטעמבער: נאך אפליקאציעס** - א ריי AI אפליקאציעס: Swifto פאר A/B טעסטן; AskMore, Merlin; Team DialMe פאר אינטערוויען; Pipio און Rephrase AI פאר ווידעאס מיט דיגיטאלע אקטיארן; Kula; GPT פאר דינגען ארבעטער; Seona פאר SEO; Snipo, AudioNotes, OneAudio; TimeOS פאר נאטיצן; HelpKit און Algomo פאר טשעטבאטס; Chat2024 מיט קלאנען פון די קאנדידאטן פאר Summit 2024; AI לעבנס-קאוטש; Generative dynamics וואס באוועגט סטילע בילדער; Trickle; AI Score My Site; ChartGenAI; CX Assist; Aragon פאר פראפעסיאנעלע פאטאס; Klu; Capsule און FireCut פאר ווידעא-באארבעטונג; Cardinal; Shortwave; Wrangle פאר זוכן ארבעט; Docera; א פירער פאר AI אין ארבעטס-פלאץ; און אן ארטיקל ווי אזוי צו באנוצן AI צו טראכטן פון ביזנעס-אידיען.

דער חודש אקטאבער 2023 איז געווען פול מיט נייע מאדעלן: Llama 2 Long, Mistral 7B ס'Meta, וואס איז אריין אין Bing Chat און ChatGPT. אין ביזנעס זענען געלט און ווערדן ווייטער געוואקסן: OpenAI האט דערגרייכט \$1.3 ביליאן יערליכע איינקונפט און פירט שמועסן איבער א ווערד פון \$80 ביליאן, Google האט צוגעזאגט ביז \$2 ביליאן פאר Anthropic, און Perplexity ווערט געשאצט אויף \$500 מיליאן. פירמעס האבן זיך גענומען צו אייגענע טשיפס, ווייל עס פעלט GPU'ס. אין דער רעגירונג-וועלט האט פרעזידענט Biden אונטערגעשריבן א גרויסע עקזעקוטיוו-ארדער איבער AI, און עס זענען צוגעקומען נייע זיכערהייט-אינישיאטיוון. אין פארשונג האט מען געזען פארשריט אין ראבאטיק, אין פארשטיין ווי מאדעלן ארבעטן אינעווייניג, און אין שנעלער אינפערענץ.

מאדעלן און נייע לאנסירונגען

- **2 אקטאבער: Meta's Llama 2 Long** – Meta האט שטילערהייט ארויסגעלאזט Llama 2 Long. ביי לאנגע פראמפטס איז דער מאדעל בעסער ווי GPT-3.5 Turbo און Claude 2.
- **2 אקטאבער: בילדער פאר ביזנעס** – דער AI טים פון Shopify האט געוויזן א פרי דעמא פון א געצייג, געבויט מיט Stable Diffusion XL, וואס בייט דעם הינטערגרונט אויף פראדוקט-בילדער. Getty Images האט צוזאמען מיט Nvidia לאנסירט "Generative AI by Getty Images". דאס געצייג שאפט בילדער פונעם ליצענצירטן פאטא-ארכיוו פון Getty.
- **3 אקטאבער: AI מכשירים צום טראגן** – דער Rewind Pendant כאפט אויף וואס מען זאגט און הערט, און שרייבט עס איבער, פארשליסט און היט עס נאר לאקאל אויפן טעלעפאן. Humane האט ארויסגעלאזט דעם "Ai Pin", א מכשיר אן א סקרין. ער האט אפטישע דערקענונג און א לעיזער-פראיעקטירטע דיספליי, און דארף נישט פארבונדן ווערן צו א סמארטפאון. שפעטער איז באריכטעט געווארן אז דער Pin ווערט באטריבן מיט GPT-4 און מען כאפט אים צו מאגנעטיש צום קלייד.
- **3 אקטאבער: Cloudflare Workers AI** – Cloudflare האט לאנסירט "Workers AI", א סערווערלאזע פלאטפארמע פאר אינפערענץ אויף איר גלאבאלער GPU נעץ. זי האט פאפולערע אפן-סאורס מאדעלן, אין צוזאמענארבעט מיט Hugging Face.
- **4 אקטאבער: Reka's Yasa-1** – די שאפער פון T5 האבן לאנסירט Yasa-1, א מאדעל פאר קול, ווידעא און טעקסט מיט שטארקע רעזולטאטן. ער איז נאר אין א פריוואטער פריוויר.
- **4 אקטאבער: LinkedIn AI געצייג** – LinkedIn האט לאנסירט AI-פונקציעס: א פארבעסערטער Recruiter, אן AI קאוטש אין LinkedIn Learning, און א געצייג פאר מארקעטינג-קאמפייןס.



LinkedIn AI געצייג · 4 אקטאבער 2023

• **5 אקטאבער: DALL-E 3** – Microsoft האט אריינגעשטעלט OpenAI's DALL-E 3 אין Bing Chat. מען פלאנט עס אויך צו ברענגען אין אנדערע פראדוקטן, ווי Paint Cocreator אין Windows. אין ChatGPT איז "Browse With Bing" ארויס פון בעטא, און DALL-E 3 איז אריין אין בעטא. OpenAI's פארשונג-ארבעט באשרייבט די הויפט-אינגרעדיענטן: א סינטעטישער קעפשאן-שאפער, בעסערע לעיטענט דיפיוזשן, און בעסערע מאסן פאר טריישאפט, סטיל און צוזאמענהאנג. א פראיעקט האט שוין נאכגעמאכט דעם "caption upsampling" מיט אפן-סאורס מאדעלן.

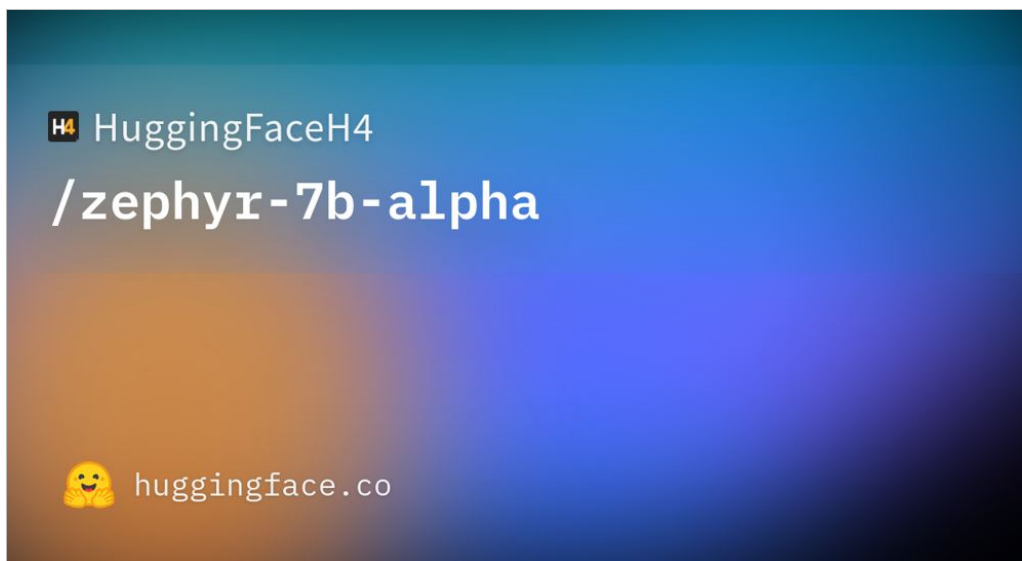
• **6 אקטאבער: AI אין בראווער און טעלעפאן** – דער Arc בראווער האט לאנסירט "Arc Max" מיט GPT-3.5 און Anthropic's מאדעלן. עס בייט נעמען פון טאבס און טעקעס, און מאכט קיצורים פון לינקס. Google's Pixel 8 און Pro 8 האבן AI פאטא-רעדאקטירונג. Google פלאנט צו באצייכענען AI-רעדאקטירטע בילדער מיט מעטאדאטא.



AI אין בראווער און טעלעפאן · 6 אקטאבער 2023

• **10 אקטאבער: Replit AI פריי** – Replit's קאוד-אסיסטענט, געטריינירט מיט Mosaic, איז יעצט פריי פאר אלע באניצער.

- **10 אקטאבער: AI פאר דאקטוירים** – Google Cloud האט לאנסירט AI-זוך געצייג פאר מעדיצינישע רעקארדס, מיט HIPAA-השגחה. Microsoft האט לאנסירט AI געצייג אין Fabric און Azure AI פאר געזונטהייט-אינסטיטוציעס.
- **11 אקטאבער: Adobe** – Photoshop איז ארויס אויפן וועב מיט Firefly AI געצייג. Adobe האט ביי MAX געוויזן Project Stardust, א פאטא-רעדאקטירער אויף Firefly Model 2. ער בייט הינטערגרונט און נעמט ארויס אביעקטן, ענליך צו Google's Magic Editor.
- **12 אקטאבער: Mistral 7B** – Mistral 7B איז באשלאגן מאדעלן צוויי מאל אזוי גרויס. די פארשונג-ארבעט באשרייבט די ארכיטעקטור, אבער נישט די טריינינג-דאטא. Hugging Face האט געטריינירט א טשעט-ווערסיע מיט direct preference optimization. דערצו זענען ארויס א Mistral מאדעל פאר מאטעמאטיק, און א Llava מאדעל אויף Mistral 7B מיט ראיה-פעאיקייטן.



Mistral 7B · 12 אקטאבער 2023

- **13 אקטאבער: Google** זוך מאכט בילדער – Google's Search Generative Experience קען יעצט שאפן בילדער פון א טעקסט-פראמפט.
- **16 אקטאבער: Eleven Labs** דאבינג – Eleven Labs האט לאנסירט א פראדוקט וואס איבערזעצט און דאבט קול אין 29 שפראכן אין סעקונדעס. עס שטיצט עטליכע רעדנער און היט אפ דעם סטיל פון יעדן רעדנער.
- **16 אקטאבער: Moonhub** רעקרוטירער – Moonhub האט לאנסירט אן AI רעקרוטירער מיט הונדערטער קונים, צווישן זיי Inflection און Anthropic. די פירמע זאגט אז מען קען דינגען 3-5 מאל שנעלער.
- **16 אקטאבער: Character.ai** גרופעס – Character.AI לאזט יעצט שמועסן אין א גרופע מיט עטליכע AI כאראקטערן אויף איינמאל.
- **18 אקטאבער: Morph** מאטעמאטיק-אסיסטענט – Morph Labs בויט אפענע געצייג פאר מאטעמאטישע באווייזן, ענליך צו Lean. דער ערשטער מאדעל, באזירט אויף Mistral, ארבעט זייער גוט.
- **18 אקטאבער: Baidu ERNIE 4.0** – Baidu האט לאנסירט ERNIE 4.0 און טענהט אז ער איז אזוי גוט ווי GPT-4.

- **19 אקטאבער:** Adept - Fuyu-8B האט ארויסגעלאזט א פשוטן מולטימאדאלן מאדעל אן אן אימעדזש-ענקאודער, וואס איז בעסער ווי 13B מאדעלן. Fuyu-8B אויף Hugging Face איז געמאכט פאר דיגיטאלע אגענטן. ער שטיצט סיי וועלכע בילד-רעזאלוציע און ענטפערט אויף גרויסע בילדער אין ווייניגער ווי 100ms.
- **24 אקטאבער:** טשעטבאטס וואקסן - Pi האט באקומען צוטריט צו נייעס און געשעענישן אויפן וועב. Claude איז יעצט צוגענגליך אין 95 לענדער.
- **30 אקטאבער:** Amazon Ads בילדער - Amazon Ads האט לאנסירט אין בעטא AI בילדער פאר אדווערטייזער. לויט באריכטן הייבט עס אויף די קליק-ראטע מיט 40%.
- **31 אקטאבער:** Open Empathic - די גרופע הינטער Stable Diffusion האט לאנסירט Open Empathic, א פראיעקט צו ברענגען דערקענונג פון געפילן אין אפן-סאורס AI.

פארשונג און טעכניק

- **2 אקטאבער:** "רעגיסטער" און "פויזע" טאקענס - Vision Transformers היטן גלאבאלע אינפארמאציע אין "אומנוציגע" פיקסלען, און דערפאר קען מען נישט פארשטיין זייערע attention maps. מיט א פשוטן [reg] טאקען ווערט עס פארריכט. אן ענליכע ארבעט האט געוויזן אז לערנבארע [pause] טאקענס פארבעסערן ריזאנירן אין שפראך-מאדעלן (ביז 1B) מיט ביז 18%.
- **2 אקטאבער:** שנעלער attention און לאנגער קאנטעקסט - Transformer-VQ גיט attention אין לינעארער צייט. StreamingLLM לאזט שטראמען טאקענס פאר טעאָרעטיש אומבאגרייכטן קאנטעקסט, אויף עקזיסטירנדע מאדעלן אן פיין-טיונינג. LightSeq טריינירט לאנג-קאנטעקסט מאדעלן בעסער ווי Megatron-LM. Simplex random feature attention נעמט אוועק די אפהענגיקייט פון דער לענג. Flash Decoding פון דער Flash Attention גרופע מאכט אינפערענץ ביז 8 מאל שנעלער.
- **2 אקטאבער:** ראיה און ווידעא - עטליכע ארבעטן: דערקענען עכטע קעגן פאלשע פנימער; ווידעא וואס פאסט צו א קול; פארבעסערן ראיה-שפראך מאדעלן נאר מיט פראמפטס; טריינירן ווידעא-מאדעלן אויף איין מאשין מיט 8 GPU'ס אין איין טאג; GAFlow פאר אפטישן פלוס.
- **3 אקטאבער:** Kosmos-2.5 - Microsoft's UniLM גרופע לייענט דאקומענטן פול מיט טעקסט פון א בילד און שאפט טעקסט אדער markdown.
- **4 אקטאבער:** אלגעמיינע ראבאט-פאליסי - Google האט געוויזן איין פאליסי מיט 55B פאראמעטערס וואס טוט איבער 160 טויזנט ארבעטן. DeepMind האט מיט 33 אינסטיטוטן געשאפן Open X-Embodiment, א דאטאבאזע מיט איבער 500 פעאיקייטן און 150,000 ארבעטן פון 22 ראבאט-טיפן.
- **4 אקטאבער:** Wayve זעלבסט-פארן - Wayve האט געוויזן א 9B מאדעל וואס פארבינדט ווידעא, אקציע און טעקסט. דערנאך איז געקומען Lingo, וואס ערקלערט זיינע באשלוסן און נעמט אקציעס.
- **4 אקטאבער:** מער ראיה-ארבעטן - מולטימאדאלע LLM'ס טרעפן בעסער אביעקטן ווען מען פארוואנדלט ערטער אין טעקסט; PC-NeRF בויט גרויסע 3D סצענעס אפילו ווען עס פעלט דאטא; DA-CLIP פאר בילד-רעסטאוראציע.
- **5 אקטאבער:** SDXL אויף TPU v5 - Google און Hugging Face ווייזן ווי שנעל SDXL לויפט אין Jax אויף די נייע TPU שפיס.

- **5 אקטאבער:** פארגעסן און טוישן וויסן – א פארשונג-ארבעט ווייזט אז טוישן וויסן אין LLM'ס קען צעמישן דעם גאנצן מאדעל. א צווייטע ארבעט "מעקט אויס" Harry Potter פון Llama 2 מיט א פראקסי-לערער מאדעל, און דער מאדעל באהאלט זיינע בענטשמארק רעזולטאטן. Salun פארבעסערט ווי AI פארגעסט דאטא.
- **5 אקטאבער:** מאדעלן אין "געזעלשאפטן" – LLM "געזעלשאפטן" מיט פארשידענע פערזענלעכקייטן פירן זיך אויף ווי מענטשן, למשל זיי גייען נאך דער מערהייט אדער דעבאטירן. AucArena ווייזט אז מאדעלן קענען גוט פלאנירן אין ליציטאציעס.
- **5 אקטאבער:** שאפן טעקסט און בילדער – MiniGPT-5 שאפט טעקסט צוזאמען מיט בילדער. MAGVIT-v2 לאזט LLM'ס באשלאגן דיפיוזשן מאדעלן ביים שאפן בילדער. א נייע טעכניק מאכט טעקסט-דיפיוזשן שנעלער.
- **6 אקטאבער:** פארשטיין נירא-נעצן – Anthropic האט גענוצט sparse autoencoders אויף א טרענספארמער מיט איין שייכט, און באקומען קלארע "פיטשערס" אנשטאט נירא-צעלן מיט עטליכע באדייטן. די פיטשערס באאיינפלוסן וואס דער מאדעל גיט ארויס. עס ווערט באצייכנט אלס א גרויסער שריט.
- **6 אקטאבער:** נאך פארשונג – GET פאר event cameras; CoDA דערקענט נייע 3D אביעקטן; Meta פארשער ליינען ווערטער פון מוח-כוואליעס אן אן אפעראציע.
- **9 אקטאבער:** מאטעמאטיק – אפן-סאורס מאדעלן לערנען זיך מאטעמאטיק מיט קאוד. AugGSM8K און MuggleMath ניצן אויגמענטירטע דאטא. ToRA ריזאנירט מיט געצייג.
- **9 אקטאבער:** בענטשמארקס – MLAGentBench לאזט AI דורכפירן ML עקספערמענטן; FACTCHD כאפט פאקטישע טעותים; א שטודיע אפשאצט שמועסן מיט פילע טורן; JudgeLM אפשאצט LLM'ס.
- **9 אקטאבער:** זיכערהייט ביי פיין-טיונינג – פארשער ווייזן אז מען קען לייכט אויסמיידן די זיכערהייט-"גרענעצן" פון Llama און GPT-3.5 Turbo מיט ביסל פיין-טיונינג. א פארשונג-ארבעט פארבונדן מיט Microsoft האט געפונען אז GPT-4 קען לייכטער מאניפולירט ווערן מיט "jailbreak" פראמפטס ווי GPT-3.5, הגם אין אלגעמיין איז ער מער פארלעסלעך.
- **9 אקטאבער:** StableLM 3B – Stability האט געטריינירט אויף 4T טאקענס צו פרובירן איבערחזון דאטא. נאכן עפנטלעכן קעגנשטעל האט מען ארויסגענומען Books3.
- **10 אקטאבער:** ראיה-ארכיטעקטורן – MSA-Conv לאזט ViTs ארבעטן מיט פארשידענע בילד-גרייסן; א נייער טרענספארמער פאר blind super-resolution; CNN'ס קענען זיין אזוי גוט ווי ViTs אין federated learning.
- **11 אקטאבער:** ראכאטן און ווידעא – מענטשלעכע באוועגונג ווערט איבערגעטראגן אויף ראכאטן מיט אן ענקאדער. FAVOR לאזט LLM'ס פארשטיין ווידעא מיט קול און בילד.
- **12 אקטאבער:** שפראך און רייד – InfoCL פארמיידט פארגעסן ביי continual learning; די סדר פון באגריפן באאיינפלוסט וויפל גוט זאצן קומען ארויס, און BART-large איז געווען דער בעסטער; קראס-מאדאלע פיוזשן פארריכט טעותים אין רייד-דערקענונג.
- **12 אקטאבער:** אוניווערסאלער סימולאטאר – DeepMind און קאלעגעס ניצן גענעראטיווע AI אלס סימולאטאר, אזוי אז ראכאט-פאליסיס זאלן ארבעטן אין דער עכטער וועלט אן עקסטרא טריינירן.

- **13 אקטאבער: Llama 2 שנעלער אנהייבן** – Anyscale האט פארקלענערט די צייט צו לאדן Llama 2 70B מיט 20 מאל, ביז בערך 30 סעקונדעס.
- **13 אקטאבער: embeddings ווייזן טעקסט** – אטאקירער קענען צוריקשאפן 92% פון טעקסטן נאר פון די embedding וועקטארס.
- **13 אקטאבער: בילד-ארבעטן** – SAM-OCTA פאר מעדיצינישע בילדער; ScaleCrafter פאר בילדער אין הויכער רעזאלוציע; "interactive text to image" מיט LLM'ס.
- **16 אקטאבער: ווייטערע פארשונג** – טרענספארמער דעקאדערס גיבן אוניווערסאלע embeddings פאר שפראך און קאוד; ECMVAE טרעפט מקורות פון קול אין ווידעא; STGM זאגט פאראויס טראפיק; א 3D מאדעל וואס מישט 3D מיט 2D.
- **17 אקטאבער: TabLiB** – דער גרעסטער עפנטלעכער טאבעלן-דאטאסעט: 627 מיליאן טאבעלן מיט 867 ביליאן טאקענס קאנטעקסט, אויף Hugging Face.
- **17 אקטאבער: ראיה און סענסארן** – שאצן ווי א קאפ איז געדרייט; SupFusion פאר LiDAR מיט קאמערע; Ex-MCR לערנט אן פארפארטע דאטא; MosaicFusion; ענדערונגען אין ערד-אבזערוואציע.
- **18 אקטאבער: LLM אנווענדונגען** – LLM4SGG פאר scene graphs; Character-LLM עמולירט היסטארישע פערזענלעכקייטן ווי Beethoven.
- **19 אקטאבער: עפעקטיווע ארכיטעקטורן** – Zipformer פאר רייד-דערקענונג איז שנעלער ווי Conformer; מען קען נאכמאכן אומסטאבילקייטן פון גרויסע מאדעלן אין קליינע מאדעלן; BitNet ניצט אסאך ווייניגער ענערגיע און זכרון.
- **20 אקטאבער: פארשטיין מאדעלן** – א טאקסאנאמיע פאר אלגעמיינערונג אין NLP; DGUA-FAS קעגן פארשטעלטע פנימער; אינטעראקטיווע גראפיקס ווי LLM'ס שטעלן פאר אמת.
- **23 אקטאבער: IBM NorthPole** – IBM'ס טשיפ, אינספירירט פונעם מוח, מאכט AI שנעלער און עפעקטיווער, ווייל ער דארף ווייניגער גיין צו דרויסנדיגן זכרון.
- **23 אקטאבער: Nvidia's Eureka** – Eureka, אן AI אגענט מיט GPT-4, טריינירט ראבאטן אליין. ער דיזינט reward funktions אויף איבער-מענטשלעכן ניווא, למשל צו דרייען א פען.
- **24 אקטאבער: קול און זען** – SALMONN פארשטייט רייד און מוזיק, און קען איבערזעצן רייד אין שפראכן וואס ער איז נישט געטריינירט אויף; א נייער מעטאד פאר lidar panoptic segmentation; BAS פאר אביעקט-לאקאליזאציע.
- **25 אקטאבער: ווידעא, מעדיצין, סימולאציע** – ווידעא-דיפיוזשן ביז 512 פריימס; SAM-Med3D, געטריינירט אויף איבער 131 טויזנט 3D מאסקעס; Meta'ס Habitat 3.0 פאר טריינירן ראבאטן.
- **26 אקטאבער: דאקומענטן און רעקאמענדאציעס** – Adobe Research ענטפערט אויף פראגן איבער לאנגע PDF'ס און האט ארויסגעגעבן א בענטשמארק; RLMRec; VidChapters-7M מיט 7 מיליאן קאפיטלען אין ווידעאס.
- **27 אקטאבער: נאך פארשונג** – GPT-4V'ס OCR פעאיקייטן; MosaicML'ס דיפיוזשן מאדעל נאר אויף Creative Commons דאטא, אויפן ניווא פון SD 2 מיט א קליינעם טייל פון דער דאטא; LLM-FP4; קוואנטיזאציע אין 4 ביטס; טיפקייט אונטער וואסער; RGM.

- **30 אקטאבער:** קליין טריינירט גרויס – NNGuide דערקענט אומבאקאנטע דאטא. Emulator Fine Tuning: מען פיין-טיונט א קליינעם מאדעל און "פראיעקטירט" עס אויף א גרויסן, און דאס ארבעט "שאקירנד גוט".
- **31 אקטאבער: GPT-3.5-turbo 20B?** – א Microsoft פארשונג-ארבעט שרייבט אז gpt-3.5-turbo האט 20B פאראמעטערס. דער ארגינעל איז געווען 175B, האט מען אפשר איבערגעטריינירט אדער דיסטילירט.
- **31 אקטאבער:** ווייטערע טעכניק – MosaicML טריינירט LLM'ס אויף AMD GPU'ס; Hugging Face האט נאכגעמאכט OpenAI'ס RLHF רעפאזיטאריע פון 2019; embeddings מיט CLIP אין ראיה; TC|IC גרופירט בילדער לויט טעקסט.

געצייג פאר פראגראמירער און אפן-סאורס

- **2 אקטאבער: Ollama** – Ollama מאכט גרינג צו לויפן און טוישן לאקאלע מאדעלן. א בלאג באשרייבט די פארבינדונג מיט LangChain.
- **2 אקטאבער:** עטליכע געצייג – SapienML – AutoML פאר פייליינס; BulkCorrector – צעטיילט טעקסט ארום דער גרענעץ פון ChatGPT; OpenCompass – אפשאצן גרויסע מאדעלן.
- **4 אקטאבער: OnnxStream** – לויפט Stable Diffusion 1.5 אויף א Raspberry Pi Zero 2.
- **5 אקטאבער: Shampoo און Claude** – Meta האט ארויסגעגעבן Distributed Shampoo אין Pytorch, אן אלטערנאטיוו צו Adam. Anthropic האט ארויסגעגעבן א Cookbook און עצות ווי צו פראמפטן Claude'ס קאנטעקסט פון 100,000 טאקענס.
- **6 אקטאבער: Ctoc און טריינירן** – Ctoc אנאליזירט קאוד אויפן טאקען-גיוואָ. א טיפער בלאג ערקלערט ZeRO, FSDP און די שוועריקייטן ביי טריינירן אויף פילע GPU'ס.
- **9 אקטאבער:** געצייג – MusicGen שטראמט מוזיק שטיקלעך; SlowLlama – פיין-טיונט Llama2-70b אויף א MacBook Air אן קוואנטזאציע; DataLens – BI; OpenAI'ס Python SDK אין בעטא.
- **10 אקטאבער: LeptonAI** – א Python פריימווערק צו בויען AI סערוויסעס.
- **12 אקטאבער:** אבזערוואציע – OpenLLMetry אויף OpenTelemetry; א סקריפט וואס פארוואנדלט ChatGPT היסטאריע אין Markdown.
- **13 אקטאבער: Microsoft Autogen** – א פריימווערק פאר AI אגענט סיסטעמען.
- **16 אקטאבער: SudoLang** – א פראגראמירונג-שפראך צו ארבעטן מיט ChatGPT, Claude, Bard און Bing Chat.
- **17 אקטאבער: AMD און נאך** – AMD האט פארעסערט ROCm שטיצע פאר RDNA GPU'ס (אריינגערעכנט RX 7900) און Libgen to txt פאר פארשן דאטא-קוואליטעט; Cal.ai – אפן-סאורס קאלענדאר; א רעפאזיטאריע איבער ChatGPT'ס סיסטעם-מעסעדזשעס.
- **18 אקטאבער: PyTorch און שנעלקייט** – PyTorch האט א לייכטן runtime פאר edge מכשירים. אן אינפערענץ ענדזשין פאר 60: diffusers שריט א סעקונדע אנשטאט 23. לעקציעס פון הונדערטער LoRA עקספערמענטן.

- **19 אקטאבער:** **Mujoco** און **MemGPT** – Mujoco 3 איז שנעלער און סטאבילער. א זאמלונג ChatGPT סיסטעם-פראמפטס. MemGPT – א לאקאלער טשעטבאט וואס לייענט און שרייבט צו זכרון און געדענקט פריערדיגע שמועסן.
- **20 אקטאבער:** **XAgent, Mojo, Waymax** – XAgent – אן אפן-סאורס אויטאנאמער אגענט; Mojo איז יעצט אויף Apple Silicon; Waymo'ס Waymax סימולאטאר אין Jax.
- **23 אקטאבער:** **קאוד און דאטא** – SolidGPT – שמועסן מיט א קאוד רעפאזיטאריע; Set of Mark prompting מיט Segment Anything; GoLLIE ציט ארויס JSON/CSV אביסל בעסער ווי gpt-3.5-turbo, און מען קען עס לויפן מיט סענסטייווע דאטא.
- **24 אקטאבער:** **פיינ-טיון דאטא** – אויטאמאטיש שאפן fine-tune.jsonl; OpenAgents – אפענע פלאגיןס.
- **25 אקטאבער:** **DeepSparse** – אינפערענץ אויף CPU מיט sparsity; Cola קאארדינירט ראיה-מאדעלן; AgentTuning.
- **26 אקטאבער:** **עטליכע געצייג** – SpellTest – AI טעסט AI; ReactAgent – React קאמפאנענטן מיט GPT-4; Polyfire – באקענד; SDXL דיסטילירט, 60% שנעלער און 50% קלענער; Promptchain.
- **27 אקטאבער:** **GPT Pilot** – שרייבט אפליקאציעס פון אנהייב לויט א באשרייבונג, שריט נאך שריט, און פרעגט ווען ער דארף הילף. דערצו א gist פאר map אויף פילע GPU'ס.
- **30 אקטאבער:** **Burn** און **AudioFlare** – Burn – דיפ לערנינג אין Rust; AudioFlare – אודיא מיט RAG Demystified; Cloudflare AI Workers.
- **31 אקטאבער:** **Insanely Fast Whisper** – 300 מינוט אודיא אין ווייניגער ווי 10 מינוט. Azure OpenAI Logger; Magnetic פאר Python.

פירמעס, געלט און ביזנעס

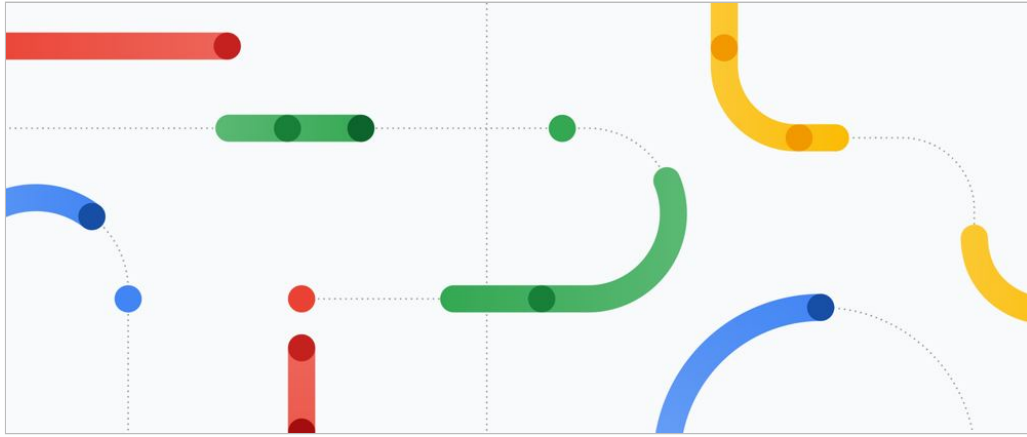
- **2 אקטאבער:** **Dataiku LLM Mesh** – א פלאטפארמע צווישן LLM צושטעלער און אפליקאציעס, מיט Snowflake, Pinecone און AI21 Labs.
- **3 אקטאבער:** **ביזנעס געצייג** – Vespio, Bestever, Collato, Strut, Semiform, Krea Logo Illusions – AI פאר פארקויף, מארקעטינג, נאטיצן, פארמס און לאגאס.
- **5 אקטאבער:** **Vespa ספין-אוף** – Yahoo האט ארויסגעצויגן Vespa אלס אן אומאפהענגיקע פירמע. Spotify און Wix ניצן עס, און עס קען באהאנדלען מיליאנען פראגעס א סעקונדע.
- **5 אקטאבער:** **Nadella אין Google פראצעס** – Satya Nadella האט געזאגט אז Google היט איר שליטה אין זוכן מיט אומיושרדיגע מיטלען, ווי די אפמאכן מיט Apple. ער האט אנגעדייט אז אפילו AI וועט שווער טוישן די געווינהייט.
- **9 אקטאבער:** **אייגענע טשיפס** – OpenAI באטראכט צו בויען אייגענע טשיפס, ווייל עס פעלט GPU'ס. Microsoft קען ווייזן איר ערשטן AI טשיפ ביי Ignite אין נאוועמבער. GitHub Copilot קאסט באניצער \$10 א חודש אבער קאסט Microsoft דורכשניטלעך \$20, און ביי טייל באניצער ביז \$80.
- **10 אקטאבער:** **OpenAI רעזידענץ** – א פראגראם פאר פארשער און אינזשעניערן פון אנדערע פעלדער, מיט \$210,000 א יאר.

- **10 אקטאבער: ChatGPT אפליקאציע** – א רעקארד פון 15.6 מיליאן דאונלאודס און בערך \$4.6 מיליאן איינקונפט אין סעפטעמבער, אבער דער וואוקס פארלאנגזאמט זיך.
- **12 אקטאבער: CleanLab \$30M** – פון אן אפן-סאורס פעקל אין MIT איז געוואקסן א געצייג אן קאוד צו גרייטן קוואליטעט-דאטא.
- **12 אקטאבער: Anthropic** – אנקלאגער ווילן נישט מען זאל דערמאנען Sam Bankman-Fried'ס \$500 מיליאן אינוועסטירונג אין Anthropic. Anthropic זוכט געלט וואס קען שטעלן איר ווערד אויף \$20-\$30 ביליאן, נאך \$1.25 ביליאן פון Amazon. Google האט צוגעזאגט ביז \$2 ביליאן. Universal Music קלאגט Anthropic, ווייל Claude גיט איבער ליד-ווערטער אן רשות.
- **13 אקטאבער: GreenBox** – Symbotic און Softbank האבן געשאפן א שותפות פאר AI לאגיסטיק אין לאגערס וואס קליינע פירמעס טיילן.
- **16 אקטאבער: Aionics און OpenAI** – Aionics ניצט AI צו געפינען עלעקטראליטן פאר באטעריעס. OpenAI'ס יערליכע איינקונפט איז \$1.3 ביליאן, קעגן \$28 מיליאן אין 2022.
- **17 אקטאבער: Nanotronics – Cubefabs** – מאדולארע טשיפ-פאבריקן מיט AI, יעדע מיט בערך 30 ארבעטער.
- **10 אקטאבער: Meta און באקאנטע** – Meta צאלט ביז \$5 מיליאן פאר זעקס שעה, צו ניצן באקאנטע פערזענלעכקייטן אלס טשעטבאטס.
- **13 אקטאבער: OpenAI פאר פראגראמירער** – OpenAI פלאנט זכרון און ראיה-פעאיקייטן אין אירע געצייג פאר פראגראמירער, וואס זאלן שטארק אראפברענגען די קאסטן. דאס ווערט ערווארט ביי דער ערשטער קאנפערענץ פאר פראגראמירער, דעם 6טן נאוועמבער.
- **19 אקטאבער: OpenAI Arrakis** – דער פראיעקט Arrakis, א מער עפעקטיווער מאדעל, איז דורכגעפאלן. OpenAI בויט א דעטעקטאר פאר AI-בילדער וואס איז ריכטיג 99% פון די מאל.
- **19 אקטאבער: Nvidia שותפות** – Nvidia מיט Foxconn בויט "AI פאבריקן" פאר אויטאנאמע קארס און ראבאטן. Nvidia און Lenovo פארברייטערן זייער שותפות פאר היבריד AI.
- **20 אקטאבער: Objective \$13M** – א סיד-רונדע פאר מולטימאדאלע זוך פאר פירמעס.
- **20 אקטאבער: Apple** – לויט באריכטן פלאנט Apple גענעראטיווע AI אין iOS 18, נישט פריער ווי שפעט 2024, אין דער וואלקן און אויפן מכשיר, מיט הונדערטער סערווערס. מען שאצט אויסגאבן פון \$4.75 ביליאן אין 2024. Daring Fireball שרייבט אז Apple בלייבט הינטערשטעליג, און עס זענען דא ספקות צווישן אינעווייניגסטע.
- **23 אקטאבער: OpenAI \$80 ביליאן** – א געשעפט אנגעפירט פון Thrive Capital וואלט געשטעלט דעם ווערט אויף \$80 ביליאן, דריי מאל מער אין ווייניגער ווי זעקס חדשים.
- **23 אקטאבער: Nvidia Jetson** – Nvidia'ס Deepu Talla האט פארגעשטעלט געצייג אין Jetson Generative AI Lab פאר ראבאטיק.
- **24 אקטאבער: Google שוץ** – Google גיט צוויי שוץ-גאראנטיעס: איינע פאר טריינירונג-דאטא, און איינע פאר רעזולטאטן, ווען מען האלט זיך צו פאראנטווארטליכע פראקטיקן.

- **24 אקטאבער: Reddit דראַעט** – Reddit קען בלאַקירן זוך-קראַולערס אויב זי באַקומט נישט באַצאלט פון AI פירמעס. איבער 535 נייעס-ארגאניזאציעס בלאַקירן שוין, און X'ס לייען-גרענעצן קומען פון AI שיקרעפינג.
- **26 אקטאבער: Perplexity** – אַ רונדע אנגעפירט פון IVP שאַצט די פירמע אויף \$500 מיליאָן. זי האט \$3 מיליאָן יערליכע איינקונפט.
- **31 אקטאבער: Snapdragon 8 Gen 3** – Qualcomm'ס טשיפּ פאר Galaxy S24 מיט AI קאמערע-געצייג.
- **16 אקטאבער: סטארטאפּ געצייג** – YC Optimizer, Softr AI, AI Content Genie, impacection.ai, Spoke, ProductAI, Helpix, AdEditor, Ween, Zaplify, ResolveAI, FlowRL, AppManager, Blaze, Nureply, Colors AI, Intently, AI SEO My Site

רעגירונגען, געזעצן און זיכערהייט

- **4 אקטאבער: NSA צענטער** – די NSA עפנט אַ צענטער פאר AI זיכערהייט.
- **9 אקטאבער: BBC** – די BBC בלאַקירט OpenAI קראַולערס, אבער זי פארמולירט פרינציפן פאר AI אין זשורנאליזם.
- **11 אקטאבער: אמעריקע קעגן EU** – וואשינגטאן ווארנט אז די EU רעגולאציעס וועלן העלפן גרויסע פירמעס, ווייל זיי פאקוסירן אויפן בויען מאדעלן אנשטאט אויף דעם ריזיקע ביים ניצן.
- **13 אקטאבער: OpenAI ווערטן** – OpenAI האט שטיל געטוישט אלע אירע "core values", מיט מער פאקוס אויף AGI.
- **25 אקטאבער: Hassabis** – Demis Hassabis רופט אז מען זאל באהאנדלען AI ריזיקעס ווי דעם קלימאט-קריזיס, מיט אן אויפזיכט-קערפערשאפט ענליך צו IPCC.
- **25 אקטאבער: Nightshade** – פארשער פון University of Chicago לאזן קינסטלער צולייגן אומזעבארע ענדערוגען, וואס פארדארבן מאדעלן וואס טריינירן דערויף. עס וועט אריין אין Glaze און ווערן אפן-סאורס.
- **26 אקטאבער: פובליק קאנסטיטוציע** – Anthropic מיט Collective Intelligence Project האבן געזאמלט פרינציפן פון בערך 1,000 אמעריקאנער דורך Polis. דער "Public" מאדעל האט מער אונטערגעשטראכן צוגעגלעכקייט און אומפארטייאשקייט.
- **27 אקטאבער: Preparedness** – OpenAI האט געשאפן א טים פאר ריזיקעס ווי סייבער און נוקלעארע קאטאסטראפּעס. Google האט פארברייטערט איר באג-באונטי פראגראם צו AI אטאקעס.
- **27 אקטאבער: Biden'ס ארדער** – Biden האט אונטערגעשריבן אן עקזעקוטיוו-ארדער מיט אכט צילן, ווי זיכערהייט, פריוואטקייט און קונה-שוץ. NIST וועט שאפן סטאנדארטן, און גרויסע מאדעל-בויער מוזן טיילן רעזולטאטן פון זיכערהייט-טעסטן.
- **30 אקטאבער: Frontier Models Forum** – Chris Meserole איז דער ערשטער דירעקטאר, און גרויסע פירמעס האבן צוגעלייגט \$10 מיליאָן צום זיכערהייט-פאנד.



2023 אקטאבער 30 · Frontier Models Forum

• **30 אקטאבער:** פראמפט-אינדזשעקשן אין בילדער – GPT-4V איז שוואך קעגן באהאלטענע אנווייזונגען אין בילדער, וואס קענען ארויסנעמען דאטא.

• **31 אקטאבער:** Hinton און Bengio – מיט 22 אנדערע עקספערטן לייגן זיי פאר אז פירמעס און רעגירונגען זאלן געבן א דריטל פונעם AI פארשונג-בודזשעט פאר זיכערהייט.

אנאליזן און מיינונגען

• **2 אקטאבער:** Altman איבער דזשאבס – לויט Sam Altman איז AI גוט ביי איינצלע ארבעטן, נישט ביי גאנצע דזשאבס, און מענטשן וועלן זיך קאנצענטרירן אויף אנדערע ארבעטן.

• **2 אקטאבער:** אלכעמיע און מוח – The Algorithmic Bridge פארגלייכט AI צו אלכעמיע, וואס איז געקומען פאר כעמיע. Exponential View פרעגט וועמען עס געהערט די דאטא פון סטארטאפס ווי Neuralink וואס ליינען דעם מוח.

• **3 אקטאבער:** SaaS און VR – Canny ברענגט אכט וועגן ווי SaaS פירמעס קענען ניצן AI. Stratechery אנאליזירט AI מיט הארדוועיר און VR. TechCrunch שרייבט אז עס איז נאך נישט קלאר וויפל קינסטלער וועלן פארדינען.

• **4 אקטאבער:** דזשאבס, דאטא, הוריקאנען – Brookings: איבער 60% פון די גענעראטיווע AI דזשאבס זענען אין 10 צענטערס. Matt Rickard פרעגט צי דאטא איז נאך א שוץ-גראבן. מאדעלן פון Nvidia, Huawei און DeepMind האבן פאראויסגעזאגט דעם הוריקאן Lee אין מינוטן, גלייך אדער בעסער ווי געווענליכע מאדעלן.

• **5 אקטאבער:** פלאטפארמע-שיפט – Digital Native פארגלייכט AI צום אינטערנעט, נישט צו סמארטפאונס. Matt Rickard זאגט אז AI וועט ווערן א באדייטנדער פלאטפארמע-שיפט.

• **6 אקטאבער:** Gemini און פויזע – Ethan Mollick (One Useful Thing) זאגט אז Gemini קען באשלאגן GPT-4. א בליק אויף א "\$68M AI hacker house". Astral Codex Ten באטראכט פארשידענע סצענארן פון א פויזע אין פארשונג. פארשער פון Microsoft און Peking University האבן גענוצט GPT-4 מיט א סאקראטישער מעטאדע אויף $NP=P$, און געקומען צו אנווייזן אז P איז נישט גלייך NP .

• **9 אקטאבער:** Benedict Evans – AI וועט מוזן ווערן ספעציפישע אנווענדונגען. אן ארטיקל אנאליזירט ביזנעס-מאדעלן פון אפן-סאורס.

- **10 אקטאבער: סטארטאפס און AMD** – וויאזוי סטארטאפס קענען בויען ווערט קעגן די גרויסע; AI איז נאר א געצייג. HPCwire: AMD קען אפשר איבערשפרינגען CUDA מיט Pytorch.
- **11 אקטאבער: טעסטן און הויט-קאליר** – פארשער ווילן איבערבויען ווי מען אפשאצט LLM'ס, אנשטאט מענטשלעכע עקזאמענס. Sony פארלאנגט א "מולטידימענסיאנעלע" מאס פאר הויט-קאליר (CIELAB), און Google און Amazon קוקן עס איבער.
- **12 אקטאבער: מולטימאדאל און פראפיטן** – Chip Huyen'ס איבערבליק פון מולטימאדאלע מאדעלן. גרויסע פירמעס קעמפן צו מאכן פראפיטן פון AI. א טענה אז GPT-4 קען ריזאנירן. Google ארבעטער פרעגן צי Bard איז כדאי.
- **13 אקטאבער: State of AI 2023** – א באריכט איבער זיכערהייט, טשיפס, און פאראויסזאגונגען איבער פילם, קאפיטאל און IPO'ס. אן ארטיקל טענהט אז OpenAI איז "צו ביליג צו באשלאגן".
- **16 אקטאבער: דזשאב-מארק** – Interconnects איבער שוועריקייטן צו געפינען דעם ריכטיגן פארשונג-דזשאב. "LLMs For Dummies" ערקלערט ווי LLM'ס ארבעטן.
- **17 אקטאבער: AGI כבר דא?** – Noema שרייבט אז היינטיגע מאדעלן קענען אנערקענט ווערן אלס ערשטע AGI. א נייע קאנפערענץ פאר שפראך-מאדעלן (COLM). Wall Street פארשטייט נישט Meta'ס אפן-סאורס וועג, אבער אין טעק זעט מען Llama ווי Noema. Linux זאגט אויך אז AI קען העלפן ארויס פון אינפלאציע, אבער ווארנט פון אומגלייכקייט.
- **18 אקטאבער: רעאל-צייט ML** – א טיפער בליק אויף רעאל-צייט ML. Ethan Mollick'ס פראגן און ענטפערס. WSJ לייענער האבן דעבאטירט צי עס מאכט אויס אויב AI האט געשאפן א קונסטווערק.
- **19 אקטאבער: טראנספארענץ** – a16z ברענגט א וועגווייזער פאר AI אין געזונטהייט. Stanford, MIT און Princeton'ס אינדעקס, מיט 100 אספעקטן, האט געפונען אז אלע פירמעס זענען נישט טראנספארענט גענוג. Llama 2 איז ערשט מיט 54%, נאך איר BloomZ און GPT-4.
- **20 אקטאבער: קליינע LLM'ס** – Matt Rickard: קליינע מאדעלן זענען שנעלער צו טריינירן, ביליגער, און לויפן לאקאל.
- **23 אקטאבער: Gary Marcus** – Gary Marcus שרייבט אז MathGLM קען נאך אלץ נישט מאטעמאטיק.
- **24 אקטאבער: זיכערהייט-דיאלאג** – אן ארגאניזירער פון AI Pause און דער יוטיובער Rob Miles שמועסן איבער AI זיכערהייט. אן ארטיקל ווארנט פון איבער-הייפ ביי סטארטאפס וואס קאפירן איינער דעם אנדערן.
- **25 אקטאבער: Gemini ליק** – לויט א ליק בייט 2 Google PaLM אויף Gemini אין Makersuite, און האט א באהאלטן געצייג "Stubbs" פאר אפליקאציע-פראטאטייפס. Seth Godin: עמאציאנעלע ארבעט אונטערשיידט אונז. Matt Rickard פרעגט צי OpenAI קען געווינען סיי קונים סיי פירמעס.
- **26 אקטאבער: געלעגנהייטן און Peak Data** – Ahmad Khan'ס Snowflake ברענגט פינף AI געלעגנהייטן. "Peak Data" באטראכט ווי אן אינטערנעט פול מיט AI דאטא קען ברענגען צו model collapse.
- **27 אקטאבער: דאטא און קאוד** – Tomasz Tunguz: דאטא קומט נענטער צום הארץ פון פראדוקט-בויען.

- **30 אקטאבער:** אליינמענט און קאפיטאלע אותיות - Guzey : עס פעלן נומערן צו מעסטן אליינמענט. OpenAI'ס פראמפטס ניצן "please" און קאפיטאלע אותיות פאר וויכטיגע אנווייזונגען.
- **31 אקטאבער:** Half-Life פון סטעק - Matt Rickard: די אינפראסטרוקטור-שיכט טוישט זיך שנעל.

ווייטערע ידיעות

- **2 אקטאבער:** אודיא סאברעדיט - א Reddit קהילה פאר AI מוזיק, רייד און קול.
- **6 אקטאבער:** Artifact - די נייעס-אפליקאציע פון Instagram'ס גרינדער שאפט יעצט AI בילדער פאר פאסטן.

דער חודש נאוועמבער 2023 איז געווען דער חודש פון OpenAI. אנגעהויבן האט זיך מיט דער ערשטער Dev Day, וואו OpenAI האט ארויסגעלאזט GPT-4 Turbo מיט 128k קאנטעקסט, די GPTs און א פלאן פאר א GPT Store. א וואך שפעטער האט דער בארד ארויסגעווארפן Sam Altman. מיקראסאפט האט אים אויפגענומען, איבער 700 ארבעטער האבן געסטראשעט אוועקצוגיין, און אין ענדע איז ער צוריקגעקומען מיט א נייעם בארד. דערנעבן האבן Anthropic לאנסירט Claude 2.1 מיט 200K טאקענס, Nvidia האט אנגעזאגט דעם H200, מיקראסאפט האט באוויזן אייגענע AI טשיפס, און Stability AI האט ארויסגעלאזט Stable Video און SDXL Turbo. אויף דער רעגירונג-זייט איז געווען Biden's עקזעקוטיוו ארדער און דער סאמיט אין Bletchley Park. Google האט אפגעשטופט Gemini ביז 2024.

מאדעלן און נייע לאנסירונגען

- **1 נאוועמבער: נייע ווערסיע פון AlphaFold – Isomorphic Labs**, א פירמע וואס איז ארויסגעוואקסן פון Google, ווייזט פאר די קומענדיגע ווערסיע פון AlphaFold. זי גייט ווייטער פון נאר פראטעינען און ארבעט בכלל בעסער.
- **1 נאוועמבער: Alibaba's Tongyi Qianwen 2.0** – Alibaba האט לאנסירט א פארבעסערטן AI מאדעל, Tongyi Qianwen 2.0. ער זאל קאנקורירן מיט Amazon און Microsoft, בעיקר אין אפליקאציעס ענליך צו ChatGPT.
- **2 נאוועמבער: Stability AI און 3D Stability AI** – האט צוערשט אנגעווינקען אויף 3D שאפונג אין obj. פארמאט און אויף אויסטוישן דעם הימל אין בילדער. דערנאך האט זי לאנסירט Stable 3D, אן אפליקאציע וואס מאכט גרינגער צו שאפן 3D מאדעלן פאר דיזיין און שפילן.
- **2 נאוועמבער: מאלטי-מאדאל Phi 1.5** – Microsoft Research האט צוגעגעבן דעם קליינעם מאדעל Phi 1.5 די פעאיגקייט צו פארשטיין בילדער, ענליך צו GPT-4, אבער מיט ווייניגער קאמפיוטער-קראפט.
- **6 נאוועמבער: xAI's ערשטער מאדעל** – Elon Musk's xAI האט טריינירט א מאדעל מיט 34 ביליאן פאראמעטערס, וואס זאל לויט די בענטשמארקס איבערשטייגן ChatGPT. ער וועט באדינען אן AI סיסטעם אינעם X פלאטפארמע. נאכדעם האט xAI ארויסגעלאזט PromptIDE, א געצייג פאר פראמפט-אינזשעניעריע און פאר פארשונג ווי מאדעלן ארבעטן.
- **6 נאוועמבער: OpenAI Dev Day** – אויף דער ערשטער דעוועלאפער-קאנפערענץ האט OpenAI באוויזן GPT-4 Turbo. ער קאסט א שטיקל פון די פריערדיגע פרייז און האט א קאנטעקסט פון 128k. נאך זענען לאנסירט געווארן "בילד-זען" אין דער API און די GPTs, אייגענע ווערסיעס פון ChatGPT וואס מען שטעלט צו אן קאוד. א GPT Store וועט בקרוב לאזן טיילן די GPTs און פארדינען דערפון, און פירמעס קענען מאכן GPTs נאר פאר זייערע אייגענע ארבעטער. Sam Altman האט געמאלדן אז ChatGPT האלט ביי 100 מיליאן באניצער א וואך און איבער צוויי מיליאן דעוועלאפער, צווישן זיי רוב Fortune 500 פירמעס. Stratechery באשרייבט די קאנפערענץ אלס א דריי צו א פראדוקט-געצילטן וועג, מיט א רמז אויף האַרדווער.

- **6 נאוועמבער: ChatGPT Plus ליינט טעקעס** – ChatGPT Plus האט באקומען נייע בעטא-פונקציעס, צווישן זיי ארויפלאדן PDF טעקעס און ארבעטן מיט זיי.
- **8 נאוועמבער: Nvidia's ChipNeMo** – Nvidia פרובירט אויס א מאדעל מיטן נאמען ChipNeMo פאר אירע טשיפ-דיזיינער. ער שרייבט סקריפטן פאר די דיזיין-געצייג, פאסט צוזאם באג-באריכטן און דינט אלס טשעטבאט מיט ערפארונג.
- **8 נאוועמבער: YouTube פרובירט AI** – YouTube טעסט נייע AI פונקציעס פאר Premium אבאנענטן. איינע לאזט פרעגן פראגעס איבער א ווידעא און באקומען פארשלאגן, און די אנדערע פאסט צוזאם די טעמעס פון די קאמענטארן.
- **9 נאוועמבער: Cohere Embed** – Cohere האט ארויסגעלאזט א נייעם embed מאדעל פאר זוכן אין אמתע, "שמוציגע" דאטא און פאר RAG סיסטעמען. ער שטיצט איבער 100 שפראכן, סיי זוכן אין דער זעלבער שפראך און סיי פון איין שפראך אין אן אנדערע.
- **10 נאוועמבער: Samsung's AI** – Samsung האט באוויזן לייוו איבערזעצונג פון טעלעפאן-שמועסן, וואס דארף ארויסקומען אין 2024. דערנאך האט זי לאנסירט Samsung Gauss, א מאדעל פאר אירע מכשירים וואס העלפט שרייבן אימעילס, אנטוויקלען ווייכווארג און ארבעטן מיט בילדער.
- **10 נאוועמבער: Adept דעמאס** – Adept בויט אן AI אגענט פאר פירמעס. זי עפנט יעצט קליינע דעמאס, למשל איבערגעבן קאפ-ארבעט, ניצן נייע געצייג און ארבעטן מיט דאטא וואס איז נישט ארגאניזירט.
- **13 נאוועמבער: Zapier און Figma** – Zapier's AI Actions לאזט יעדע AI פלאטפארמע אויספירן איבער 20,000 אויטאמאטיש-אקציעס דורך באפעלן אין פשוטער שפראך. Figma האט אריינגעשטעלט AI אין FigJam, איר דיגיטאלע טאוול, כדי מען זאל בעסער קענען צוזאמענארבעטן.
- **14 נאוועמבער: GPT-5** – OpenAI ארבעט אויף GPT-5 און זוכט נאך געלט פון Microsoft כדי אים אנצוגיין.
- **14 נאוועמבער: Nvidia H200** – Nvidia לאנסירט דעם HGX H200 טשיפ. ער האט אסאך מער מעמארי און מעמארי-בענדווייט ווי דער H100, און זאל ארויסקומען אין 2024, כאטש מען ווייסט נאך נישט ווען ער וועט טאקע צו באקומען זיין.
- **15 נאוועמבער: GPT-4 Turbo אויף Azure** – Microsoft וועט אנבאטן GPT-4 Turbo אויף Azure OpenAI Service ביזן סוף יאר, ביליגער און מיט מער קאנטראל. OpenAI's Custom Models פראגראם וועט ארבעטן צוזאמען מיט Microsoft's סיסטעם.
- **16 נאוועמבער: מיקראסאפט'ס טשיפס** – Microsoft האט אנטוויקלט אן אייגענעם AI טשיפ, Azure Maia, און אן Arm-באזירטן פראצעסאר, Azure Cobalt, פאר אירע דאטא צענטערס. דער ציל איז ווייניגער אנגעוויזן צו זיין אויף Nvidia. ביידע קומען ארויס אין 2024.
- **16 נאוועמבער: Google SGE גלאבאל** – Google's "Search Generative Experience" איז געעפנט געווארן אין 120 נאך לענדער און שפראכן. Google גייט פאמעלעך, ווייל עס רירט אן אירע הויפט-זוך-סיסטעמען.
- **16 נאוועמבער: GitHub Copilot** – Copilot Chat, באזירט אויף GPT-4 און וואס לאזט פראגראמירן אין פשוטער שפראך, קומט אין דעצעמבער Copilot Enterprise 2023, צוגעפאסט צום קאוד פון א פירמע, קומט אין פעברואר 2024.

- **17 נאוועמבער: Lyria מיט YouTube – DeepMind** האט אנגעזאגט א נייע מוזיק-סיסטעם, Lyria. א גרויסער טייל פונעם פארבעסערונג איז געקומען פון א דאטא-שותפות וואו די קינסטלער ווערן באצאלט. מען וועט דאס קענען ניצן דורך YouTube Shorts. יעדע טראק באקומט אן אומהערבארן SynthID סימן, וואס בלייבט אפילו נאך קאמפרעסיע.
- **17 נאוועמבער: Meta's ווידעא-עדיטירן – Meta** האט אנגעזאגט מאדעלן צו שאפן און עדיטירן ווידעא. דער חידוש איז אז מען באהאנדלט יעדע שאפונג אלס אן אנווייזונג, און דערפון קומט ארויס די מעגליכקייט צו עדיטירן. די ארכיטעקטור איז אויך פשוטער געווארן.
- **17 נאוועמבער: Azure אוואטארן – Microsoft** האט לאנסירט Azure AI Speech text-to-speech avatar. מען קען שאפן אוואטארן וואס זעען אויס ווי אמת און זאגן א געשריבענעם טעקסט אין פארשידענע שפראכן.
- **21 נאוועמבער: Bing Chat ווערט Copilot – Microsoft** רופט יעצט Bing Chat מיטן נאמען "Copilot", כדי עס זאל זיין איין נאמען איבער אלע אירע פראדוקטן. דאס אנטהאלט אויך Bing Chat Enterprise.
- **22 נאוועמבער: Windows AI Studio – Microsoft** האט אנגעזאגט Windows AI Studio פאר דעוועלאפער, מיט געצייג פון Azure AI Studio און Nvidia Hugging Face. האט דערהיינטיגט TensorRT-LLM פאר RTX 30/40 קארטלעך, כדי מען זאל קענען לויפן מאדעלן לאקאל אויפן קאמפיוטער.
- **22 נאוועמבער: Gemini אפגעשטופט – Google** האט אפגעשטופט Gemini ביזן ערשטן קווארטאל פון 2024. מען האט מורא געהאט צי ער וועט קענען גלייכן זיך מיט GPT-4, און מען וויל קודם פארשטארקן די פראדוקטן פאר געווענליכע באניצער.
- **22 נאוועמבער: Stable Video Diffusion – Stability AI** האט ארויסגעלאזט Stable Video Diffusion, אן אפן-סאורס מאדעל וואס שאפט 14 בילדער-ראמען ווידעא. ער שליסט זיך אן צו איר סעריע מאדעלן פאר בילד, שפראך, קול, 3D און קאוד.
- **22 נאוועמבער: ChatGPT מיט קול – OpenAI** האט עס ארויסגעלאזט מיט א ווייז איבער דער מעשה מיט איר CEO. באניצער וואס צאלן נישט. OpenAI האט עס ארויסגעלאזט מיט א ווייז איבער דער מעשה מיט איר CEO.
- **22 נאוועמבער: Amazon Olympus – Amazon** לויט באריכטן טריינירט Amazon א מאדעל מיטן נאמען Olympus מיט 2 טריליאן פאראמעטערס, צוויי מאל אזויפיל ווי GPT-4. דער ציל איז איבערצושטייגן GPT-4 און Bard.
- **27 נאוועמבער: Claude 2.1 – Anthropic** האט ארויסגעלאזט Claude 2.1 מיט א קאנטעקסט פון 200K טאקענס, און ער זאגט האלב אזוי אפט פאלשע זאכן. ער האט א בעטא פונקציע צו ניצן געצייג און APIs, און קען אויך זוכן אויפן וועב. צוזאמען דערמיט האט Anthropic אראפגעלאזט די פרייזן, אונטער דרוק פון סיי אפענע און סיי פארמאכטע מאדעלן.
- **28 נאוועמבער: Tesla FSD v12 – Tesla** גיט די Full Self-Driving v12 נאר פאר אירע ארבעטער. דער הויפט-חידוש איז אז דאס קאנטראלירן דעם קאר טוען נורא-נעצן אנשטאט קאוד וואס אינזשעניערן האבן געשריבן. עס קען דערגרייכן צו קונים ביזן סוף יאר. Elon Musk האט פריער געזאגט אז ער גלייבט אז Tesla וועט דערגרייכן אמת'ע זעלבסט-פארן ביזן סוף יאר.

- **29 נאוועמבער:** CICERO – CICERO Meta'ס איז דער ערשטער AI אגענט וואס האט דערגרייכט א מענטשליכן ניוואו אינעם סטראטעגיע-שפיל Diplomacy.
- **29 נאוועמבער:** SDXL Turbo – Stability AI'ס SDXL Turbo שאפט בילדער מיט הויכער רעזאלוציע אין 207 מיליסקונדעס. ער ניצט נישט LCM, נאר א נייע טעכניק וואס הייסט Adversarial Diffusion Distillation. די וועגן (weights) און דער קאוד זענען צו באקומען פאר פארשונג.

פארשונג און טעכניק

- **1 נאוועמבער:** פארשטיין ווידעא – MPVSS צעטיילט ווידעא נאר אין שליסל-ראמען און זאגט פאראויס די מאסקעס פאר די איבעריגע. TESTA פארגיכערט פארשטיין לאנגע ווידעאס דורך צוזאמשטעלן ענליכע ראמען, מיט ווייניגער רעכענען און בעסערע רעזולטאטן.
- **1 נאוועמבער:** GPT-4V אין מעדיצין – א פארשונג פרובירט ווי גוט GPT-4V ענטפערט פראגעס איבער מעדיצינישע בילדער, ווי X-rays און CT.
- **1 נאוועמבער:** נאך פארשונג – א מעטאד מיט א "restoration gap" פאררעכט פלענער פון דיפיוזשאן מאדעלן וואס זענען נישט ריאליסטיש. COMM פאראייניגט די שטארקייטן פון CLIP און DINO אין מאלטי-מאדאלע מאדעלן. PUCA איז א נייער וועג צו רייניגן בילדער פון "ראש", וואס היט אפ J-invariance.
- **2 נאוועמבער:** ספארס טענזארס – פארשער פון MIT און NVIDIA האבן אנטוויקלט טעכניקן צו געפינען גיך נולן אין נירא-נעצן און אריינפאסן מער ווערטן וואס זענען נישט נול אויפן טשיפ. דאס פארגרעסערט די שנעלקייט און שפארט ענערגיע.
- **2 נאוועמבער:** לערנען פון טעותים – LeMa לערנט מאדעלן חשבון דורך לערנען פון טעותים, ווי א תלמיד. א נירא-נעץ ארויסגעגעבן אין Nature קען אלגעמיינערן שפראך ענליך צו מענטשן, און איז בעסער ווי ChatGPT.
- **3 נאוועמבער:** לאנגער קאנטעקסט פאר Mistral – Nous Research האבן מיט YaRN אויסגעשטעלט Mistral 7B צו ארבעטן מיט ביז 128k טאקענס, כמעט אן פארלוסט ביי קורצע טעקסטן. Amazon האט ארויסגעלאזט א Mistral 7B מיט 32k טאקענס.
- **3 נאוועמבער:** עטליכע פארשונגען – א דאטא-סעט מאכט מאדעלן מער מיטפילנדיג פאר פסיכאלאגישע עצות. מאדעלן ארדענען קלאסן אין א הייערארכיע, וואס פארבעסערט בילד-קלאסיפיקאציע ביי CLIP. Cohere For AI האט ארויסגעלאזט קאוד פאר PEFT ביי Mixture-of-Experts מאדעלן.
- **6 נאוועמבער:** NeurIPS, HelixNet, סקיצעס – Paper Digest האט צוזאמגעפאסט פארשונג-ארבעטן פון NeurIPS 2023. HelixNet שטעלט צוזאם דריי Mistral מאדעלן: איינער שאפט, איינער קריטיקירט און איינער שרייבט איבער, און דאס פארבעסערט די רעזולטאטן שטארק. RT-Sketch לאזט א ראבאט באקומען אנווייזונגען דורך א גראבע סקיצע.
- **7 נאוועמבער:** RedPajama v2 – RedPajama v2 איז א דאטא-סעט פון 30 טריליאן טאקענס, 2.5 מאל גרעסער ווי וואס מען האט גענוצט פאר GPT-4. עס האט נישט קיין קאוד, אבער עס האט פינף שפראכן, און די דופליקאטן זענען שוין ארויסגענומען.

- **7 נאוועמבער: בילדער און רעקאמענדאציעס** – Gaussian Mixture Solvers מאכן דיפיוזשאן בילדער שארפער אן צו פארלירן גיכקייט. DADA פאראייניגט 3D מיט RGB צו געפינען פעלערן אויף אויבערפלאכן. CLLM4Rec פאראייניגט מאדעלן מיט ID סיסטעמען פאר רעקאמענדאציעס.
- **8 נאוועמבער: SCT, PixArt, DARE** – SCT לאזט קליינע מאדעלן שאפן embeddings ווי גרויסע. PixArt, א טעקסט-צו-בילד מאדעל מיט א diffusion transformer, קאסט 90% ווייניגער צו טריינרן ווי UNet מאדעלן. DARE לאזט צולייגן נייע פונקציעס צו מאדעלן ווי BERT אן איבערטריינרן.
- **9 נאוועמבער: אויטאס, מוזיק, יאפאניש** – RobustMat פארבעסערט ווי זעלבסט-פארנדיגע קארס דערקענען לאנדמארקס. MusicGen שאפט יעצט סטעריא, וואס פארבעסערט די ערפארונג כאטש נישט די בענטשמארקס. JaSPICE מעסט די קוואליטעט פון יאפאנישע בילד-באשרייבונגען.
- **10 נאוועמבער: טריינרן אויף ריזיגע פארנעמען** – א פארשונג ווייזט ווי צו רייניגן טריינינג-דאטא. Google האט טריינירט אויף איבער 50 טויזנט טשיפס צוגלייך מיט איבער 66% אויסנוץ. MLPerf v3.1 ווייזט אז דער H100 איז דער שטארקסטער, אבער לויט פרייז האלט זיך Intel גוט.
- **13 נאוועמבער: אטאקעס און 3D** – DG-SCT איז א נייער attention מעכאניזם פאר קול-און-בילד מאדעלן. א בלאג-ארטיקל צייכנט אויס די סארטן אטאקעס קעגן שפראך-מאדעלן און ווי מען שיצט זיך. 3DStyle-Diffusion באצירט 3D מעשעס מיט 2D דיפיוזשאן.
- **14 נאוועמבער: RoseTTAFold All-Atom** – א נירא-נעץ וואס מאדעלירט מאלעקולן אויסער פראטעינען: מעדיצין, וויטאמינען, DNA און צוקער. מען האט מיט דעם דיזיינט נייע מאלעקולן וואס ווערן באשטעטיגט אין לאבאראטאריע.
- **14 נאוועמבער: FinGPT, VLPMarker** – FinGPT איז טריינירט אויף פיניש. דאס איבערחזר'ן דאטא מיט עטליכע טריקס האט געגעבן גלאטע רעזולטאטן, וואס קען העלפן פאר שפראכן מיט ווייניג טעקסט אויפן אינטערנעט. VLPMarker שיצט vision-language מאדעלן מיט א וואסער-סימן.
- **15 נאוועמבער: Capybara, דעקאנטאמינאציע** – Nous Research האבן ארויסגעלאזט Capybara, א 34B מאדעל אויף Yi, טריינירט אויף לאנגע שמועסן. LMSYS ווייזט אז ווען מען פארמולירט איבער טעסט-דאטא, קען מען פארפאלשן בענטשמארק רעזולטאטן. MonoDiffusion שאצט אפ טיפקייט אן אמת'ע דאטא.
- **15 נאוועמבער: באהאלטענע חשבונות** – לויט Redwood Research קענען מאדעלן באהאלטן זייערע צווישן-שריטן אינעם טעקסט אויף א וועג וואס מענטשן קענען נישט לייענען.
- **16 נאוועמבער: ברילן, DPO, וועטער** – RISC-V טשיפ מיט א קליינעם YOLO מאדעל דערקענט זאכן אויף קלוגע ברילן מיט ווייניג ענערגיע. DPO פארבעסערט ווי ריכטיג מאדעלן זענען, און רעדוצירט מעדיצינישע טעותים מיט איבער DeepMind 50%. האט ארויסגעגעבן קאוד און וועגן פון איר 10-טאג וועטער מאדעל.
- **17 נאוועמבער: VideoQA און נאך** – Flipped-VQA פארבעסערט ענטפערס איבער ווידעא דורכדעם וואס דער מאדעל קוקט מער אויפן ווידעא אליין. א דאטא-סעט דערקענט ווי תלמידים פירן זיך אין קלאס. SentAlign ציילט צוזאם זאצן אין גרויסע פאראלעלע דאקומענטן.

- **20 נאוועמבער: פארשידענע** – TopoMLP פארשטייט וועגן פאר זעלבסט-פארנדיגע קארס. א פארשונג טראגט איבער וויסן פון איין מאדעל צום אנדערן מיט ווייניג קאמפיוטער-קראפט. א מעמארי-מאדול פארבעסערט דערקענען זאכן אין פארשידענע סביבות.
- **21 נאוועמבער: Video LLaVa, Orca 2** – Video LLaVa ארבעט מיט ווידעא און דערגרייכט די בעסטע רעזולטאטן אין צוזאמפאסן ווידעא. Orca 2's Microsoft, צו באקומען אויף Hugging Face, איז בעסער ווי גרעסערע מאדעלן אין לאגיק. א פארשונג ווייזט אז פארמאכטע מאדעלן האבן ווארשיינליך נישט טריינירט אויף די טעסטן.
- **22 נאוועמבער: SDXL און Lookahead** – א ארטיקל ערקלערט דעם latent space פון SDXL. Lookahead Decoding פון LMSYS פארגיכערט ענטפערס ביז 2.3 מאל, אן א צווייטן draft מאדעל.
- **22 נאוועמבער: GraphCast** – לויט א פארשונג אין Science איז DeepMind's GraphCast בעסער ווי די טראדיציאנעלע וועטער פאראויסזאגונגען, ביז 10 טעג פאראויס.
- **27 נאוועמבער: ZipLoRA, SegVol** – ZipLoRA פאראייניגט LoRAs און היט אפ דאס נייע פון יעדן איינעם. SegVol צעטיילט פילע אנאטאמישע טיילן אין CT בילדער. Visual In-Context Prompting פארבעסערט צעטיילן בילדער.
- **27 נאוועמבער: GenSLMs** – א מאדעל פאר גענעטישע דאטא, טריינירט נאר אויף פרי-פאנדעמיע דאטא, האט געשאפן סעקווענצן ענליך צו שפעטערע COVID ווערסיעס.
- **28 נאוועמבער: Starling-7B** – Berkeley האט טריינירט Starling-7B מיט סינטעטישע דאטא פון וואס מען בעוואַרצוגט (RLAIF), און ארויסגעלאזט סיי דעם מאדעל און סיי דעם reward מאדעל. אנדערע ארבעטן: 360-גראד פאנאראמעס, און בילדער פארוואנדלט אין געמעלן.
- **28 נאוועמבער: Monosemanticity** – פארשער ביי Anthropic האבן געפונען א וועג צו פארשטיין די אינעווייניגסטע ארבעט פון א מאדעל. זיי ווייזן אז באזונדערע באגריפן ווערן ארויסגעשטעלט אין באזונדערע features.
- **29 נאוועמבער: דיפיוזשאן** – StableVSR פארבעסערט ווידעא רעזאלוציע. DiffSLVA מאכט אנאנים צייכן-שפראך ווידעאס און היט אפ דעם אינהאלט. א נייע מעטאד שאפט קלענערע און מער פארשידנארטיגע דאטא-סעטס.
- **29 נאוועמבער: AlgoPerf** – MLCommons' AlgoPerf איז א קאנקורענץ מיט געלט-פרייזן פאר וועגן צו טריינירן גיכער, למשל גיכער ווי ADAM.

געצייג פאר פראגראמירער און אפן-סאורס

- **2 נאוועמבער: LangChain, Llama, Mistral** – LangChain Templates גיבן מוסטערן פאר אפליקאציעס. Meta גיט א וועגווייזער ווי צו אנהייבן ארבעטן מיט Llama. א notebook ווייזט ווי צו fine-tune Mistral 7B פאר \$1.
- **2 נאוועמבער: SAM און MetaCLIP** – מען צעטיילט בילדער לויט טעקסט מיט Segment Anything און MetaCLIP.
- **3 נאוועמבער: RAG געצייג** – א פיטאן געצייג פאר RAG, fine-tuning און סערווירן מאדעלן; Talk Llama לאזט רעדן מיט LLaMA אין טערמינאל.

- **6 נאוועמבער: פיר געצייג** – Video2Music שאפט מוזיק וואס פאסט צום געפיל פון א ווידעא. Qdrant'ס vector database אין Rust באדינט X'ס "ענליכע פאסטס". Giskard טעסט מאדעלן אויף פעלערן ווי פארגעלונג (bias), אויסרינען פון דאטא און הלוצינאציעס. Deepspeed'ס נייע ווייכווארג צו לויפן מאדעלן איז ביז 2.4 מאל גיכער.
- **7 נאוועמבער: TopicGPT, FACTOR** – TopicGPT געפינט טעמעס אין טעקסט. FACTOR דערקענט פאלשע (deepfake) מעדיע אן טריינירן. OpenAI האט ארויסגעלאזט א consistency decoder פאר Stable Diffusion אלס אפן-סאורס.
- **8 נאוועמבער: CogVLM** – CogVLM-17B (10B פאר זען, 7B פאר שפראך) איז אן אפענער מאדעל וואס גייט שטארק אויף בענטשמארקס. OVIR-3D זוכט 3D זאכן לויט טעקסט. Langroid מאכט AI אגענטן. א notebook באשרייבט ספארט-געשעענישן מיט OpenAI געצייג.
- **9 נאוועמבער: RoboFlow, Punica** – RoboFlow ווייזט GPT-4V צוזאמען מיט SAM און DINO. Punica לויפט פילע LoRA מאדעלן פארן פרייז פון איינעם. I2VGen-XL שאפט ווידעא אין צוויי שטאפלען.
- **10 נאוועמבער: LongLoRA און נאך** – LongLoRA טאפלט דעם קאנטעקסט. lidar העלפט דערקענען וועגן מיט ווייניגער האנט-ארבעט. טייל פון טעקסט-מאדעלן ארבעטן גוט פאר בילד-אויפגאבן. VimGPT בלעטערט דאס וועב מיט GPT-4V.
- **13 נאוועמבער: Zephyr און נאך** – Hugging Face'ס alignment handbook, נאך Zephyr. Tarsier גיט געצייג פאר וועב-אגענטן. LLaVA-Plus ניצט געצייג צו עדיטירן און שאפן בילדער. איין מאדעל פארבעסערט די קוואליטעט פון אלערליי קול. Bark ניצט מען פאר שנעלע לאנגע רעדע. AI town לויפט לאקאל מיט RWKV.
- **14 נאוועמבער: OpenGPTs** – אן אפן-סאורס אלטערנאטיוו צו GPTs, באזירט אויף LLaMA. LangChain. מיט 16,000 טאמיל טאקענס. BestGPTs, א רשימה פון די בעסטע GPTs.
- **15 נאוועמבער: AMBER, Instructor** – AMBER מעסט הלוצינאציעס אין מאלטי-מאדאלע מאדעלן. KnowPAT פארבעסערט ענטפערס אין ספעציפישע פעלדער. "Draw a UI" מאכט HTML פון א סקיצע. Instructor ציט ארויס סטרוקטורירטע דאטא מיט OpenAI function calling.
- **16 נאוועמבער: Qwen Audio** – Alibaba'ס מאדעל פאר קול-צו-שפראך. "distance indexing" פארבעסערט ווי מען פילט אויס ראמען צווישן ווידעא-בילדער. Mistral 7B לויפט אויף א Raspberry Pi. LLM Sherpa איז אן API.
- **17 נאוועמבער: SAM 8x גיכער** – Segment Anything איז 8 מאל גיכער מיט PyTorch טריקס. דערצו א וועגווייזער פאר מאדעלן אין פראדוקציע, און א רשימה פון AI פעלערן (exploits) וואס זענען ריכטיג געמאלדן געווארן.
- **20 נאוועמבער: StyleTTS 2** – StyleTTS 2 צילט צו מענטשליכער רעדע. דערצו קומען Neural-Cherche פאר זוך-מאדעלן, Tensorli אין Numpy, און מקורות פאר דאטא אינזשעניערן.
- **21 נאוועמבער: Make it Real** – tldraw ציכנט א דיזיין און GPT-V מאכט פון אים ארבעטנדיגן קאוד. CoWPiRec פארבעסערט רעקאמענדאציעס. RustGPT איז א ChatGPT קלאן אין Rust.
- **22 נאוועמבער: LangChain איז Go** – LangChain ארבעט יעצט אין Go, אויך מיט Ollama. מען קען אויפשטעלן א GGUF מאדעל אין 3 מינוט, און עס איז דא א וועגווייזער פאר פראדוקציע.

- **27 נאוועמבער: Duck Track – Duck Track, DPO** – דא א ביישפיל-אימפלעמענטאציע פון DPO, און Kandinsky Video, טעקסט-צו-ווידעא.
- **28 נאוועמבער: Fast Whisper** – מען שרייבט אראפ 2.5 שעה קול אין 100 סעקונדעס, און מען דערקענט ווער רעדט. Agency איז פאר Go. CoachLM פארבסערט אנווייזונג-דאטא.
- **29 נאוועמבער: Tanuki – Tanuki, Set of Mark** – מאכט אפליקאציעס וואס ווערן ביליגער מיט דער צייט. LLaVa. Animatable Gaussians vision מאדעלן, אויך LLaVa. Animatable Gaussians. שאפט מענטשן-אוואטארן פון ווידעא.
- **1 נאוועמבער: קליינע אפליקאציעס** – Phot.AI (פאטא-עדיטירן), MagicPost (LinkedIn), Keyword Generator (SEO), Julius (דאטא אנאליז), Layer (ספארט 40% פון פארשונג-צייט), Salesforce (אימעילס), Orba, CopyRemix, Success.ai, OHMYSYNT, Hiring Tool, VocAI (סופפארט-פראגעס) און Tactiq (נאטיצן מיט GPT-4 Turbo). Lindy בויט "AI ארבעטער" אן קאוד אין 30 סעקונדעס. Reverb פילט אויס פארמס אן הלוצינאציעס.
- **28 נאוועמבער: Align API – Align API** – כאפט אויף סענסיטיווע פראמפטן איידער זיי גייען צום מאדעל, און שפארט אזוי קאסטן.

פירמעס, געלט און ביזנעס

- **2 נאוועמבער: AWS – AWS Capacity Blocks** – לאזט רעזערווירן Nvidia GPUs פאר א באשטימטע צייט.
- **3 נאוועמבער: Meta און אפן-סאורס – Llama 2** – לייסענס טרעפט נישט אלע באדינגונגען פון OSI. Meta's VP פאר פארשונג רעדט וועגן באלאנסירן טיילן וויסן מיט ביזנעס-קאסטן.
- **3 נאוועמבער: Apple – Tim Cook** – זאגט אז Apple אינוועסטירט "גאנץ שטארק" אין AI, אבער ער האט נישט ערקלערט ווי.
- **7 נאוועמבער: Xbox און Inworld** – Xbox ארבעט צוזאמען מיט Inworld AI אויף א copilot פאר שפיל-דיזייןער און א מאטאר פאר AI פיגורן.
- **8 נאוועמבער: Luminance** – א בריטישע פירמע האט אנטוויקלט AI וואס פארהאנדלט אליין קאנטראקטן.
- **9 נאוועמבער: Google's אינוועסטירונגען** – Google האט פארברייטערט איר שותפות מיט Anthropic, מיט TPU v5e טשיפס. זי איז אויך אין שמועסן צו אינוועסטירן הונדערטער מיליאנען אין Character.AI.
- **9 נאוועמבער: Copilot איבער מיליאן** – GitHub Copilot האט איבער 1 מיליאן באצאלטע באניצער אין 37,000 ארגאניזאציעס, א וואקס פון 40% אינעם קווארטאל.
- **9 נאוועמבער: Hugging Face H4** – די H4 גרופע אנטוויקלט אפן-סאורס טשעטבאטס און alignment טעכניקן.
- **10 נאוועמבער: OpenAI דאטא-שותפות** – OpenAI זוכט שותפים צו שאפן דאטא-סעטס פאר טריינירן, למשל די רעגירונג פון איסלאנד און Free Law Project. TLDR האט באמערקט אז דאס קען ווייזן אז OpenAI לויפט אויס טאקענס.

- **13 נאוועמבער: Poe – Poe** לאזט שעפער פון באטס פארדינען דורך טיילן אבאנעמענט-געלט און צאלן פער מעסעדזש. דערווייל נאר אין אמעריקע.
- **14 נאוועמבער: AGI ביים בארד – OpenAI** ס'בארד פון 6 מיטגלידער וועט באשליסן ווען מען האט דערגרייכט AGI, וואס קען ווירקן אויף דער פארבינדונג מיט Microsoft.
- **15 נאוועמבער: Retool באריכט – Retool** האט געפרעגט 1,500 ארבעטער: MongoDB Atlas איז די פאפולערסטע vector database. די שוועריגקייטן זענען ווי ריכטיג די ענטפערס זענען און דאטא-זיכערהייט.
- **15 נאוועמבער: You.com APIs – You.com** ס'APIs לאזן מאדעלן צוקומען צו פרישע אינפארמאציע פון וועב, אנגעהויבן פון \$100 א חודש.
- **15 נאוועמבער: IBM – IBM** האט אוועקגעלייגט \$500 מיליאן פאר א פאנד צו אינוועסטירן אין AI סטארטאפס פאר פירמעס.
- **15 נאוועמבער: Humane AI Pin** – א מכשיר אן א סקרין פאר \$699 און \$24 א חודש, אויף T-Mobile. עס טוט איבערזעצן, פאטאגראפירן און ארבעט אלס פערזענליכער געהילף.
- **20 נאוועמבער: OpenAI קרייז – OpenAI** דער בארד האט ארויסגעווארפן Sam Altman ווייל ער האט "נישט מער קיין צוטרוי". Greg Brockman און אנדערע פארשער זענען אוועקגעגאנגען. Emmett Shear איז געווארן צייטווייליגער CEO, און Microsoft האט אויפגענומען Altman און Brockman. איבער 700 ארבעטער האבן געסטראשעט אוועקצוגיין צו Microsoft, וואס פארמאגט 49% און האט צוגעזאגט פלעצער פאר אלע. Salesforce האט אנגעבאטן פולע באצאלונג פאר OpenAI ארבעטער, און Shear האט געסטראשעט אוועקצוגיין אויב דער בארד ווייזט נישט קיין ראיה. אין ענדע זענען Altman און Brockman צוריקגעקומען מיט א נייעם בארד און א פלאן פאר אן אויספארשונג.
- **20 נאוועמבער: OpenAI פארמאכט Plus – OpenAI** האט אפגעשטעלט נייע אבאנעמענטן פאר ChatGPT Plus צוליב דעם גרויסן פארלאנג.
- **20 נאוועמבער: Discord – Discord** פארמאכט איר טשעטבאט Clyde.
- **21 נאוועמבער: Amazon קורסן – Amazon** גיט אומזיסטע קורסן אין generative AI.
- **21 נאוועמבער: BCV Labs – Bain Capital Ventures** עפנט אן AI אינקובאטאר מיט אפיס, GPU קרעדיטן און געלט.
- **27 נאוועמבער: Codegen – Codegen** האט אויפגעשטעלט \$16 מיליאן פאר אנגענטן וואס שרייבן קאוד, פון Jira ביז GitHub.
- **27 נאוועמבער: Kyutai** – א נייע פארשונג-לאבאראטאריע אין פאריז, א נאן-פראפיט, האט אויפגעשטעלט \$330 מיליאן.
- **28 נאוועמבער: Amazon און Salesforce** – זיי האבן פארברייטערט זייער שותפות פאר דאטא און AI.
- **29 נאוועמבער: Casetext – Casetext** איז פארקויפט געווארן צו Thomson Reuters פאר \$650 מיליאן, נאך א לאנגן וועג.
- **29 נאוועמבער: נישט קיין בארד-זייץ פאר OpenAI – Microsoft** וועט ווארשיינליך נישט געבן קיין זייץ אויפן נייעם בארד פאר Microsoft און אנדערע אינוועסטארן.

רעגירונגען, געזעצן און זיכערהייט

- **1 נאוועמבער:** קאפירייט – א ריכטער האט אוועקגעווארפן רוב טענות פון קינסטלער קעגן Midjourney און DeviantArt. די Verge האט צוזאמגעזאמלט די טענות פון AI פירמעס פארוואס זיי דארפן נישט צאלן פאר אינהאלט מיט קאפירייט.
- **1 נאוועמבער:** Sutskever – Ilya Sutskever האט זיך אריבערגעדרייט צו "superalignment", פארהיטן אז א superintelligence זאל זיך נישט פירן שלעכט. ער גלייבט אז AI וועט בקרוב איבערשטייגן מענטשן.
- **1 נאוועמבער:** פיישינג – AI וועט שטארק פארבעסערן פיישינג, וואס ווערט שוין גענוצט אין 90% פון סייבער אטאקעס.
- **2 נאוועמבער:** Biden און Blatchley – Biden's עקזעקוטיוו ארדער האט נישט קיין "צייך", אבער עס ווייזט די ריכטונג, מיט ווירקונגען אויף סטארטאפס. אויפן סאמיט אין Blatchley Park האבן לענדער אונטערגעשריבן א דעקלאראציע איבער AI זיכערהייט, און דער אמעריקאנער האנדל-דעפארטמענט האט אנגעזאגט אן אינסטיטוט פאר AI זיכערהייט.
- **3 נאוועמבער:** Isambard-AI – בריטאניע אינוועסטירט \$273 מיליאן אין א סופער-קאמפיוטער וואס די אוניווערסיטעט פון Bristol בויט.
- **6 נאוועמבער:** זיכערהייט – Google פארברייטערט איר VRP פראגראם צו generative AI, און שיצט דעם "supply chain" פון AI מיט SAIF. א פאסט באשרייבט אטאקעס וואס פארגיפטן דעם דאטא-פייפליין. א פארשער האט ארויסגעצויגן פראמפטן פון אנדערע באניצער דורך א Google Doc, און Google האט עס פאררעכט.
- **14 נאוועמבער:** Civitai – Civitai האט אריינגעשטעלט "bounties" פאר AI בילדער פון אמת'ע מענטשן, א פונקציע וואס מען קריטיקירט.
- **17 נאוועמבער:** YouTube סימנים – שעפער אויף YouTube מוזן מעלדן ווען זיי ניצן AI אין ווידעאס.
- **20 נאוועמבער:** Meta צעלאזט גרופע – Meta האט צעלאזט איר Responsible AI גרופע. רוב מיטגלידער זענען אריבער צו generative AI.
- **27 נאוועמבער:** Stability רעזיגנאציע – א פירער ביים קול-טיים פון Stability AI האט רעזיגנירט. ער האלט אז טריינירן אויף קאפירייט אינהאלט אן רשות איז נישט "fair use".
- **29 נאוועמבער:** Sports Illustrated – Sports Illustrated האט בסוד ארויסגעגעבן ארטיקלען פון פאלשע AI שרייבער.

אנאליזן און מיינונגען

- **1 נאוועמבער:** פחד און רעגולאציע – Andrew Ng זאגט אז גרויסע פירמעס פארגרעסערן דעם פחד פון AI כדי צו פארשטיקן קאנקורענץ און אפן-סאורס. The Algorithmic Bridge האלט אז פחד פירט צו געאיילטע געזעצן. Big Technology שרייבט אז די "דומערס" באקומען יעצט קעגנדרוק.
- **1 נאוועמבער:** ביזנעס – ארטיקלען איבער Gen Z'ס השפעה אויפן מארק, און פארוואס "flywheels" פון פירמעס ווערן שוואכער.

- **2 נאוועמבער: AI אין ביונעס** – Paul Adams פון Intercom רעדט איבער AI און פראדוקט-סטראטעגיע. AI ענדערט מארקעטינג אויף 4 וועגן. AI וועט פארגרעסערן די מאסע אינהאלט. AI אין ווייכווארג פירמעס.
- **6 נאוועמבער: Cruise** – Gary Marcus פרעגט צי Cruise איז "דער Theranos פון AI", ווייל זייערע קארס דארפן אפט מענטשן פון דער ווייטנס.
- **7 נאוועמבער: סטארטאפס און GPT** – דאס לייענען PDFs אין ChatGPT האט אויסגעמעקט א ריי סטארטאפס, א ווארענונג פאר פירמעס וואס בויען אויף גרויסע פירמעס. Insider Intelligence זאגט פאראויס אז 25% פון אמעריקאנער אינטערנעט-באניצער וועלן ניצן AI יעדן חודש.
- **8 נאוועמבער: GPTs, פאראויסזאגונגען** – GPT-4 איז צו זיכער אין זיינע פאראויסזאגונגען (5000 פראגעס פון Manifold). Big Technology צייכנט אויס די שוואכקייטן פון OpenAI. Ethan Mollick שרייבט אז GPTs זענען כמעט אגענטן.
- **9 נאוועמבער: AI קעגן קריפטא** – אן אנאליז פון 2 מיליאן קאמענטארן אויף Hacker News ווייזט אז AI איז קעסיידער פאפולער.
- **10 נאוועמבער: Nvidia און Gates** – איין ארטיקל באשרייבט ווי Nvidia איז געווארן דער "קעניג-מאכער" אין Bill Gates 2023. זאגט פאראויס אז AI אגענטן וועלן זיין די גרעסטע ענדערונג אין קאמפיוטערס זינט די גראפישע אינטערפעיס.
- **14 נאוועמבער: רעוויי** – Sebastian Raschka איבער AI אין Angular Ventures 2023. טיילט אויף AI סטארטאפס אין דריי גרופעס.
- **16 נאוועמבער: מוזיק און Copilot** – a16z האלט אז AI מוזיק קען זיין א "Midjourney מאמענט". Matt Rickard זאגט אז Copilot שטארקט דווקא די עקזיסטירנדע ביונעסער.
- **17 נאוועמבער: דעמאקראטיע** – א קריטיק אויף OpenAI'ס רוף פאר דעמאקראטישע פראצעסן אין AI.
- **20 נאוועמבער: Coatue** – Sri Viswanath'ס 117-זייטיגער באריכט: אפן-סאורס איז דאס הארץ פון AI, און דאס בעסטע שטייט נאך פאר אונז. דאס אינטערעס פון דעוועלאפער איז געוואקסן און דערנאך געפאלן.
- **20 נאוועמבער: אנאליז פון OpenAI** – ארטיקלען פרעגן צי OpenAI איז נאך א נאן-פראפיט. Tomasz Tunguz האלט אז Microsoft וועט געווינען, און אז פירמעס ווי Anthropic קענען פראפיטירן פון דער קריזיס.
- **21 נאוועמבער: SaaS ווערט** – AGI קען דריקן אויף דעם ווערט פון SaaS פירמעס.
- **27 נאוועמבער: Q*** – א ליק זאגט אז די קריזיס ביי OpenAI איז ארויסגעקומען פון א דורכברוך וואס פאראייניגט זוכן מיט RL. Interconnects ברענגט א טעאריע וואס דאס קען זיין.
- **28 נאוועמבער: ארגאניזאציעס** – Mollick שרייבט ווי מען דארף איבערבויען ארגאניזאציעס פאר AI. Casper Labs האט געפרעגט 600 IT מנהלים איבער בלאקטשעין און AI.
- **29 נאוועמבער: קאסטן** – Tomasz Tunguz זאגט אז מאדעלן זענען דערווייל באגרענעצט דורך די קאסטן, אבער דאס קען זיך ענדערן.

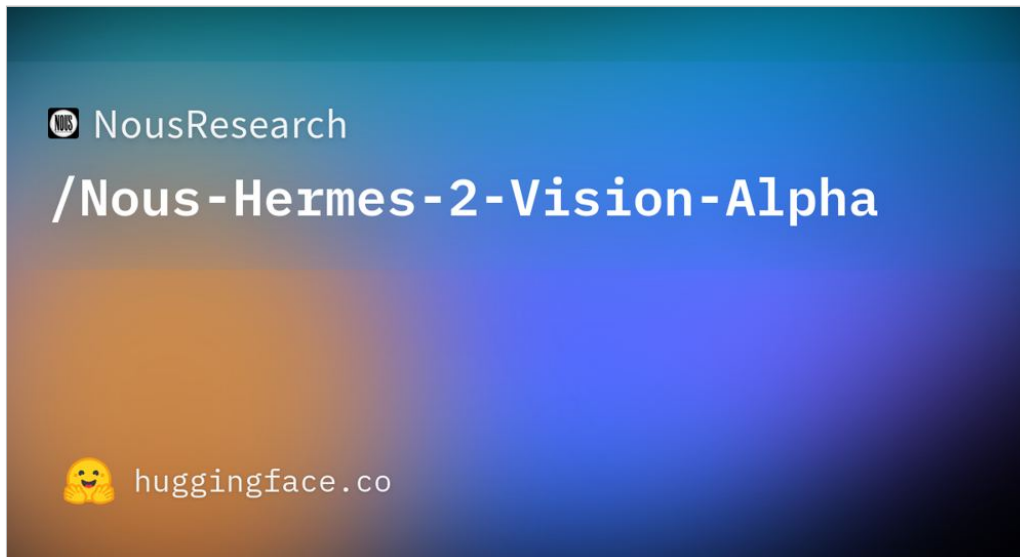
ווייטערע ידיעות

- **3 נאוועמבער: Beatles** – די Beatles האבן ארויסגעלאזט "Now and Then", מיט John Lennon'ס קול אויפגעלעבט דורך AI.
- **16 נאוועמבער: DeepMind קונסט** – DeepMind האט ארויסגעגעבן אנימאציעס וואס העלפן פארשטיין AI.

דעם חודש איז געווען פול מיט גרויסע מאדעל-לאנסירונגען: Google האט ארויסגעלאזט Gemini און AlphaCode 2, אבער איר בעסטע דעמא-ווידעא איז קריטיקירט געווארן אלס פארפאלשט. Mistral האט ארויסגעלאזט א Mixture of Experts מאדעל, א פלאטפארמע מיט API, און איז גרייט צו אויפנעמען איבער 400 מיליאן דאלאר. די נאכווייען פון Sam Altman'ס אפזאגן און צוריקקומען צו OpenAI זענען ווייטער געגאנגען מיט נייע באריכטן איבער די סיבות. די אייראפעאישע יוניאן האט דערגרייכט א פרעלימינארן אפמאך אויף דעם ערשטן פולשטענדיגן AI געזעץ. אויך האבן זיך באוויזן קליינע אבער שטארקע מאדעלן ווי Phi-2, און נייע ארכיטעקטורן וואס זענען נישט באזירט אויף טראנספארמערס. דעם חודש (פון 18טן ביז 22סטן דעצעמבער) האט Google ארויסגעלאזט Gemini Pro אויף Vertex AI און פובליקירט די גרויסע Gemini פארשונג-ארבעט. אויפן טשיפ-פראנט האבן AMD און Intel געפרוואווט צו פארמעסטן זיך מיט Anthropic. Nvidia איז געווען אין די ידיעות מיט נייע באדינגונגען און שמועסן איבער \$750 מיליאן ביי א ווערט פון \$15 ביליאן. ByteDance איז באשולדיגט געווארן אז זי האט געניצט OpenAI'ס טעכנאלאגיע אקעגן די כללים. דערצו זענען ארויס א נייע פארשונג-ארבעטן איבער וויזשן, 3D, רעדע-צו-טעקסט און שנעלערע מאדעלן.

מאדעלן און נייע לאנסירונגען

- **4 דעצעמבער: Amazon Q** האט פראבלעמען – Amazon'ס נייער טשעטבאט Amazon Q ליידט לויט באריכטן פון "האלוצינאציעס" (אויסגעטראכטע ענטפערס). ער לאזט אויך ארויס סענסיטיווע אינפארמאציע, צווישן אנדערע איבער AWS דאטא צענטערס.
- **4 דעצעמבער: ScienceGPT** מיט א טריליאן פאראמעטערס – די אמעריקאנער רעגירונג און Intel האבן אנגעהויבן טריינירן ScienceGPT, א מאדעל מיט איין טריליאן פאראמעטערס.
- **5 דעצעמבער: Hermes 2.5** וויזשאן מאדעל – Nous Research האט ארויסגעלאזט אן אפן-סאורס מאדעל פאר טעקסט און בילדער. ער איז געבויט אויף דעם בעסטן 7B מאדעל מיט SigLIP, און קען לויפן אויף געווענליכע קאמפיוטערס. א נייעס דערביי איז "function calling", ד.ה. דער מאדעל קען רופן פונקציעס אין קאד. ער איז נאך אין אלפא-שטאפל צוליב א פראבלעם מיט האלוצינאציעס.



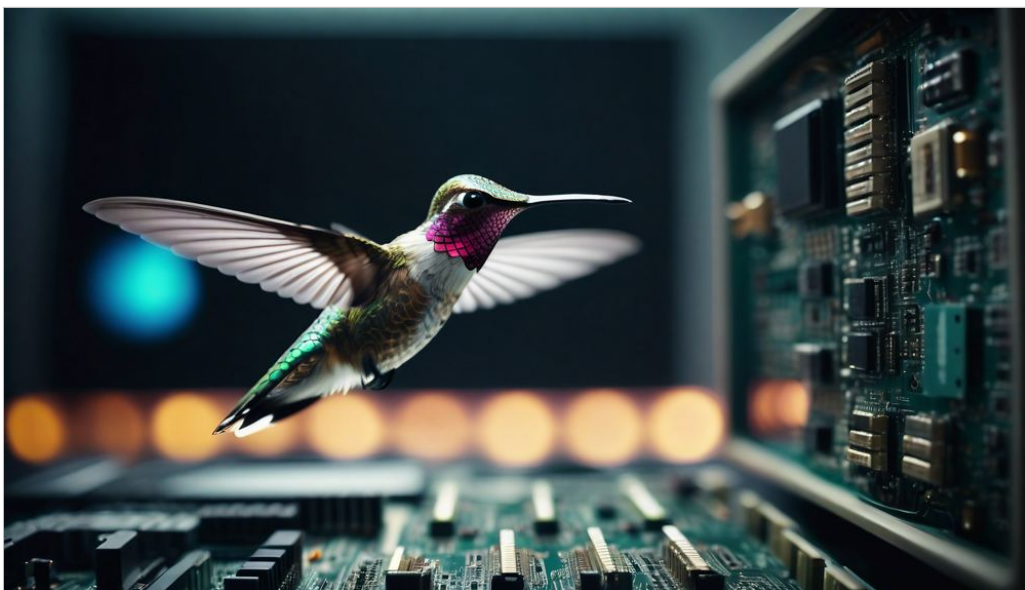
Hermes 2.5 וויזשאן מאדעל · 5 דעצעמבער 2023

- **5 דעצעמבער: Perplexity's "אנליין" מאדעלן** – Perplexity AI האט ארויסגעלאזט צוגעפאסטע ווערסיעס פון אפן-סאורס מאדעלן פון Mistral און Meta. זיי זענען די ערשטע לייוו API'ס וואס שטיצן זיך אויף אינפארמאציע פון וועב-זוכן, כדי צו געבן פאקטישע און דערהיינטיגטע ענטפערס. לויט VentureBeat קען דאס אנטראנען Google Search.
- **6 דעצעמבער: Google's מוזיק-עקספערימענט** – Instrument Playground ניצט MusicLM צו שאפן מוזיק-שטיקלעך פון 20 סעקונדעס, אינספירירט פון איבער 100 אינסטרומענטן פון איבער דער וועלט. מען קען זיי ווייטער צופאסן אין א "sequencer" און זיי אראפלאדן.
- **7 דעצעמבער: Google Gemini און AlphaCode 2** – Google's הויפט-מאדעל Gemini איז פון אנהייב געבויט פאר אלע סארט אינפארמאציע (טעקסט, בילדער און מער), און אין טעקסט איז ער גלייך מיט GPT-4. אויף דעם יסוד האט DeepMind טריינירט AlphaCode 2, וואס איז בעסער ווי 85% פון מענטשליכע פראגראמירער אין פארמעסטן אויף Codeforces. AlphaCode 2 וועט זיין צוגעגליך אין Google's API נאך דעם 13טן דעצעמבער. אבער Google's דעמא-ווידעא איז קריטיקירט געווארן: די אינטעראקציעס זענען נישט געווען לייוו, נאר צוגעגרייט מיט סטילע בילדער און טעקסט. אן ערשטע אנאליז פון Roboflow ווייזט אז Gemini איז בעסער ווי אסאך אפענע מאדעלן אין בילד-אויפגאבן, און זייער קאנקורענטפֿעאיג מיט GPT-V.



2023 7 דעצעמבער · AlphaCode 2 און Google Gemini

- **7 דעצעמבער: Meta's Imagine** – Meta האט לאַנסירט "Imagine", א בילד-שאפער וואס מאכט בילדער פון טעקסט, אויף אן אייגענעם וועבזייטל.
- **7 דעצעמבער: Copilot באקומט נייע מאדעלן** – Microsoft לייגט צו צו Copilot דעם GPT-4 Turbo, א פארבעסערטן DALL-E 3, א "code interpreter" (א געצייג וואס פירט אויס קאוד), און "Deep Search" אין Bing.
- **8 דעצעמבער: Grok** אויף X – Grok, דער קאנקורענט צו ChatGPT פון Elon Musk'ס סטארטאפ xAI, איז לאַנסירט געווארן אויף X פאר Premium Plus אבאנענטן אין אמעריקע.
- **8 דעצעמבער: Kandinsky 3.0** – א טעקסט-צו-בילד מאדעל וואס שאפט נאך מער רעאליסטישע בילדער.
- **11 דעצעמבער: StableLM Zephyr 3B** – Stability AI האט ארויסגעלאזט א טשעט-מאדעל מיט 3 ביליאן פאראמעטערס, אינספירירט פון Zephyr 7B. ער איז צוגעפאסט מיט דער מעטאד (Direct Preference Optimization) כדי צו פאלגן אנווייזונגען און ענטפערן אויף פראגעס.

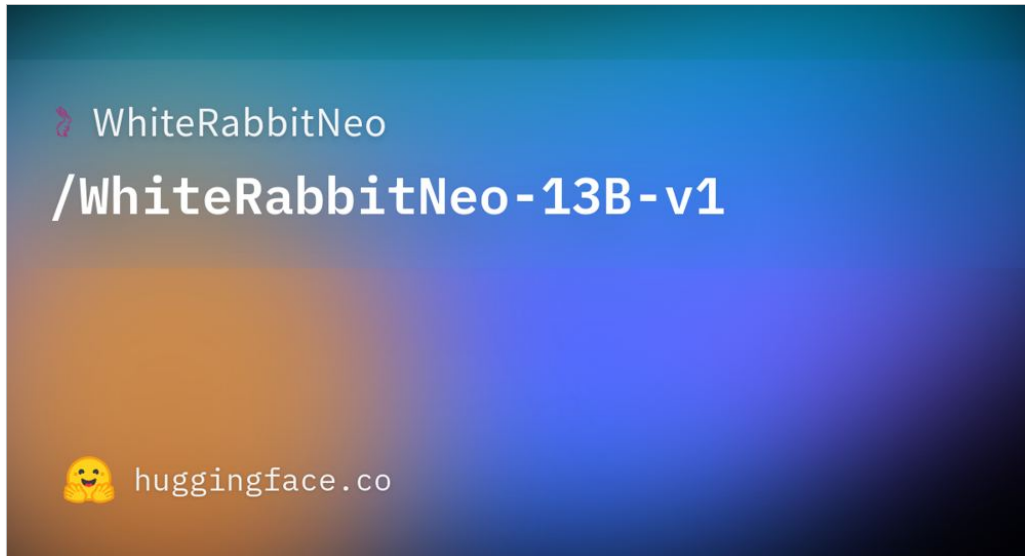


2023 11 דעצעמבער · StableLM Zephyr 3B

- **11 דעצעמבער: Mistral:** נייער מאדעל, API און געלט – דער 8 חדשים אלטער סטארטאפ Mistral גרייט זיך אויפצונעמען איבער 400 מיליאן דאלאר. ער האט ארויסגעלאזט אן 8x7B "Mixture of Experts" מאדעל, וואס איז צוזאמגעשטעלט פון עטליכע ספעציאליזירטע טיילן. Mistral האט אויך געעפנט אן API מיט נייע מאדעלן, אן "embedding" מאדעל, און מאדעלן צוגעפאסט צו פאלגן אנווייזונגען. אומארויסגעלאזטער מיטל-גרויסער מאדעל איז שוין בעסער ווי GPT-3.5.
- **12 דעצעמבער: Meta's אודיא-מאדעל – Meta** האט פארעפנטלעכט די פארשונג-ארבעט און דעמאס פון א יסוד-מאדעל פאר אודיא, וואס קען שאפן קאנטראלירבארן קלאנג אין פארשידענע סטילן, אלעס פון איין מאדעל.
- **13 דעצעמבער: Phi-2 – Microsoft's Phi-2** האט 2.7 ביליאן פאראמעטערס און איז טריינירט מערסטנס אויף "סינטעטישע" דאטא (וואס איז געשאפן געווארן דורך קאמפיוטער). ער איז גלייך מיט מאדעלן פון 13 ביליאן פאראמעטערס. די שוועריקייט איז צו דערקענען צי טייל פון די טעסט-פראגעס זענען אריינגעפאלן אין די טריינירונג-דאטא אין אן איבערגעשריבענער פארעם. ער איז צוגענגליך אין Azure, און זיין "weights" (די אויסגעלערנטע ציפערן פונעם מאדעל) זענען ארויף אויף Hugging Face פאר פארשונג.
- **14 דעצעמבער: Imagen 2 – Google** האט לאנסירט Imagen 2 פאר באשטעטיגטע באניצער פון Vertex AI. ער קען שרייבן טעקסט אין בילדער, שאפן לאגאס, און איבערזעצן אין פארשידענע שפראכן.
- **15 דעצעמבער: Capybara און Obsidian – Nous Research** האט ארויסגעלאזט Capybara, א שפראך-מאדעל, און Obsidian, זיין ווערסיע פאר בילדער. ער האט 3 ביליאן פאראמעטערס און קען לויפן אויף טעלעפאנען.
- **15 דעצעמבער: Samsung Gauss – Samsung's AI** מאדעל ווערט שוין גענוצט צו פארבעסערן די פראדוקטיוויטעט פון אירע ארבעטער, און וועט שפעטער אריינקומען אין אירע פראדוקטן. ער האט מאדולן פאר שפראך, הילף מיט קאוד, און בילד-רעדאקטירן.
- **18 דעצעמבער: Gemini Pro אויף Vertex AI – Google** האט לאנסירט Gemini Pro, א לייכטערע ווערסיע פון איר פארגעשריטענעם מאדעל Gemini Ultra. ער איז צוגעשטעלט אלס עפנטלעכע פרובע-ווערסיע אין Vertex AI. דער מאדעל באהאנדלט טעקסט און בילדער, און מען קען אים צופאסן פאר פארשידענע באנוצן.
- **18 דעצעמבער: AI בילדער אין Snapchat+ – Snapchat** פירט איין נייע AI מעגלעכקייטן פאר Snapchat+ אבאנענטן. עס איז דא א בילד-מאכער פון טעקסט, א "Dream" סעלפי וואס שאפט פאנטאזיע-בילדער מיט פריינט, און אן "extend" געצייג וואס פארלענגערט אויטאמאטיש דעם הינטערגרונט פון א פאטא. די מעגלעכקייטן ווערן ארויסגעלאזט געגנט נאך געגנט פאר איבער 7 מיליאן אבאנענטן.
- **19 דעצעמבער: נייע Intel טשיפס – Intel** האט ארויסגעלאזט די 5טע דור Xeon פראצעסארן מיט א פאקוס אויף AI פאר קלאוד, נעץ און "edge" באנוצן. דאס איז א טייל פון Intel's סטראטעגיע זיך צו פארמעסטן אויפן AI טשיפ מארק, וואו Nvidia הערשט.
- **19 דעצעמבער: Atlassian Intelligence – Atlassian** האט לאנסירט אירע ערשטע AI סערוויסעס איבער איר קלאוד פלאטפארמע, באזירט אויף 20 יאר דאטא. צווישן די הויפט-מעגלעכקייטן: גענעראטיווע AI אין רעדאקטאר, AI צוזאמענפאסונגען און אויטאמאציע אין נאטירלעכער שפראך. דער

ציל איז אויטאמאטיזירן רוטינע ארבעט, ערקלערן פירמע-אויסדרוקן און העלפן ביי באשלוסן באזירט אויף דאטא.

- **20 דעצעמבער:** סייבער-זיכערהייט מאדעל – א נייע פירמע, whiterabbitneo (א רמז אויף "Matrix"), האט ארויסגעלאזט א מאדעל מיט 13 ביליאן פאראמעטער. ער איז טריינירט פאר סייבער-זיכערהייט, סיי פאר אטאקע סיי פאר פארטיידיקונג. די ידיעה רופט די פירמע "קאנטראווערסיאל".



סייבער-זיכערהייט מאדעל - 20 דעצעמבער 2023

- **20 דעצעמבער:** Google's VideoPoet – VideoPoet איז א שפראך-מאדעל פון Google פאר שאפן ווידעא. ער טוט פארשידענע ארבעטן ווי טעקסט-צו-ווידעא, ענדערן דעם סטיל פון א ווידעא, און אפילו ווידעא-צו-קלאנג. דאס נייע איז אז ער פאראייניגט עטליכע מעטאדן פון ווידעא-שאפונג אין איין מאדעל.
- **20 דעצעמבער:** Copilot קאמפאנירט לידער – Microsoft Copilot קען יעצט שאפן לידער, דורך א פארבינדונג מיט דער מוזיק-אפליקאציע Suno.
- **21 דעצעמבער:** Zoo טעקסט-צו-CAD – Zoo Dev (פריער Kitty Cad) האט זיך איבערגערופן און ארויסגעלאזט א טעקסט-צו-CAD מעגלעכקייט. זי שאפט 3D אביעקטן וואס מען קען דרוקן אדער ניצן אלס טיילן.

פארשונג און טעכניק

- **4 דעצעמבער:** בילדער, אנימאציע און "דיפוזיע" – א "דיפוזיע מאדעל" שאפט בילדער דורך ביסלעכווייז אויסרייניגן א "ראש" (ראנדאמאלע פונקטן). MoMask שאפט אנימאציע פון טעקסט מיט פייער קאנטראל. א נייע מעטאד ניצט StyleGAN כדי בעסער צו האלטן דאס געזיכט פון א מענטש ענלעך אין געשאפענע בילדער. אנדערע ארבעטן שאפט דיפוזיע מאדעלן אן "attention" (דער הויפט-מעכאניזם פון טראנספארמערס), און ניצט אנשטאט דעם "state space" מאדעלן, א נייע סארט ארכיטעקטור וואס ווערט אינטערעסאנט. אין טעאריע ברענגט דאס מער שנעלקייט.
- **4 דעצעמבער:** עטליכע פארשונג-פראיעקטן – MaXTron – אן איינהייטלעכע ארכיטעקטור צו צעטיילן ווידעאס אין זייערע טיילן. א פשוטע מעטאד צו אנהייבן א קליינעם מאדעל פון א גרויסן, וואס פארבעסערט

- דראמאטיש די ארבעט פון קליינע מאדעלן. SAFE – פארבינדט בילדער, "event streams" (דאטא פון ספעציעלע קאמערעס) און באצייכענונגען צו דערקענען מוסטערן.
- **5 דעצעמבער: "Linear-time" מאדעלן** – א נייע state space ארבעט נעמט אידייען פון LSTM (אן עלטערע סארט נעץ). א מאדעל מיט 3 ביליאן פאראמעטערס איז בעסער ווי טראנספארמערס מיט 7 ביליאן.
 - **5 דעצעמבער: דיפויזע-מעטאדן** – MotionEditor – רעדאקטירט באוועגונג אין ווידעאס און היט אפ דעם הינטערגרונט און דאס אויסזען פון דער הויפט-פערזאן. IMMA – באשיצט טעקסט-צו-בילד מאדעלן כדי מען זאל זיי נישט קענען צופאסן צו שאפן שעדליכע אינהאלט. DMP – ניצט אויסגעלערנטע דיפויזע מאדעלן פאר אויפגאבן ווי אפשאצן 3D אייגנשאפטן, מיט ווייניג טריינירונג-דאטא.
 - **6 דעצעמבער: 3D, בילדער און ווידעא** – SplaTAM – ניצט "Gaussian Splatting", א מעטאד צו בויען 3D סצענעס, צו נאכפאלגן אן קאמערע און COLMAP (א געווענליך געצייג פאר דעם). די רעזולטאטן זענען שטארק, אבער עס פאדערט נאך שווערע חשבונות. DiffiT – "vision transformers" פאר בילדער אין הויכע רעזאלוציע. Zero123++ – שאפט עטליכע זייטן פון א זאך פון איין בילד. RDVS און DCTNet+ – נייע דאטא און א נעץ צו דערקענען וויכטיגע זאכן אין ווידעאס מיט טיפקייט-אינפארמאציע. "Style" – האלט א גלייכן סטיל צווישן פארשידענע בילדער, אן קיין נאכטריינירן.
 - **6 דעצעמבער: אינסטרוקציע-טריינירונג און צופאסן צו א פאך** – א טיפע איבערבליק איבער נייע פארשריטן אין "instruction tuning" (לערנען א מאדעל צו פאלגן אנווייזונגען), ווי למשל סינטעטישע דאטא. Morph ניצט סינטעטישע דאטא וואס דער מאדעל שאפט אליין, כדי צו לערנען ספעציעלע ידיעות. דאס איז בעסער ווי "fine-tuning" (נאכטריינירן דעם מאדעל) און ווי RAG (צוגעבן דאקומענטן בשעת'ן ענטפערן).
 - **7 דעצעמבער: בענטשמארקס און פארשונג** – BenchLMM – א בענטשמארק צו פרובירן ווי שטארק מאדעלן ווי GPT-4V און LLaVA האלטן זיך ביי פארשידענע בילד-סטילן. א פארשונג-ארבעט באשרייבט מיט ווערטער דעם חילוק צווישן צוויי גרופעס בילדער. Diffusion-SS3D – דערקענט זאכן אין 3D מיט ווייניג באצייכנטע דאטא. אן ארבעט ניצט דעם "Oogiri" הומאר-שפיל צו פרובירן ווי מאדעלן קענען טראכטן שעפערש.
 - **8 דעצעמבער: פראמפט פאר לאנגע קאנטעקסט אין Claude 2.1** – Claude 2.1 קען נעמען 200 טויזנט "tokens" (שטיקלעך טעקסט) אויף אמאל. אין דרויסנדיגע טעסטן האט ער אבער געטראפן נאר 27% פון אינפארמאציע וואס איז באהאלטן אין מיטן פון לאנגע טעקסטן. ווען מען לייגט צו צום פראמפט דעם זאץ "Assistant: Here is the most relevant sentence in the context", שפרינגט ער ארויף צו 98%.
 - **8 דעצעמבער: OneLLM און Free3D** – OneLLM פארשטייט אכט סארטן דאטא, צווישן זיי בילדער, אודיא, און אפילו מוח-אקטיוויטעט, דורך פארבינדן אלעס צו שפראך. Free3D שאפט 3D אנבליקן פון איין בילד, אן א שווערן 3D מאדעל.
 - **11 דעצעמבער: נייע מעטאדן און ארכיטעקטורן** – RAVE – רעדאקטירט ווידעאס מיט טעקסט-צו-בילד מאדעלן און היט אפ די באוועגונג. KTO – א מעטאד צו "align" (צופאסן צו מענטשליכע ווינטשן) א מאדעל. עס פאדערט נאר א "יא" אדער "ניין" פאר יעדע ביישפיל, אנשטאט טייערע פארגלייכן "A" איז בעסער ווי "B". StripedHyena 7B – א state space מאדעל וואס ווייזט אז נייע ארכיטעקטורן קענען

ברענגען לענגערע קאנטעקסט. א מאדעל פאר שאפן סצענעס, וואס פארבינדט טעקסט, באוועגונג און עקזיסטירנדע זאכן.

• **12 דעצעמבער: ביאלאגיע, ווידעא און RL – BioCLIP** – א מאדעל פאר בילדער אין ביאלאגיע, בעסער מיט כמעט 20% ווי OpenAI's CLIP אויף ביאלאגישע אויפגאבן. מיט אים איז ארויס א זאמלונג פון 10 מיליאן בילדער מיט טעקסט. STDN – קלאסיפיצירט ווידעאס. FamO2O – פארבעסערט "reinforcement learning" (לערנען דורך באליונג). א שטודיע איבער די שוועריקייטן פון צעטיילן זאכן אין בילדער און באצייכנטע דאטא.

• **13 דעצעמבער: 3D און בילדער – CorresNeRF** – פארבעסערט NeRF (א מעטאד צו בויען 3D סצענעס פון בילדער) ווען מען האט ווייניג אנבליקן. DiAD – געפינט אומנארמאלע זאכן מיט דיפוזיע. TGP-T – רעדוצירט שטארק דעם GPU זכרון פאר מאדעלן ווי CLIP. AGAP – רעדאקטירט 3D סצענעס דורך א 2D בילד. Sherpa3D – שאפט 3D זאכן פון טעקסט מיט גוטע קוואליטעט און ריכטיגע פארעם.

• **14 דעצעמבער: עפעקטיווקייט און מעדיצין – FIND** – אן אינטערפעס וואס לאזט מאדעלן בעסער פארשטיין בילדער און טוישן דעם מאדעל אליין. א מעטאד צו קאמפרעסירן שפראך-מאדעלן אן טייערע ייזנווארג. HyperRouter – מאכט עפעקטיווער דאס טריינירן. אן איבערבליק איבער CLIP אין מעדיצינישע בילדער.

• **15 דעצעמבער: FunSearch געפינט נייע מאטעמאטיק** – DeepMind האט פארעפנטלעכט אין Nature א פארשונג-ארבעט איבער FunSearch. עס ניצט א שפראך-מאדעל צו זוכן פונקציעס וואס לייען אפן פראבלעמען אין מאטעמאטיק און קאמפיוטער-וויסנשאפט. ביים "capset" פראבלעם האט עס געפונען טייל פון די גרעסטע capsets ביז היינט.

• **15 דעצעמבער: OpenAI's "סופער-אליינמענט"** – OpenAI's סופער-אליינמענט גרופע ארבעט אויף ווי צו קאנטראלירן AI וואס איז קליגער ווי מענטשן. זי האט באוויזן אז א שוואכער מאדעל קען פירן א פיל שטארקערן: מיט GPT-2 האבן זיי צוריקגעקראגן א גרויסן טייל פון GPT-4's אליינמענט-ארבעט, אזוי ווי שוואכע מענטשן וועלן מוזן פירן סופער-אינטעליגענטע מאדעלן. לויט זיי האט די מעטאד א קלארן וועג צו שטארק פארבעסערט ווערן, נישט ווי RLHF (טריינירן מיט מענטשליכע אפשאצונגען). OpenAI טיילט אויך אויס גרענטס פאר סטודענטן און פארשער אין דעם פעלד; מען דארף זיך איינגעבן ביז 18טן פעברואר.

• **15 דעצעמבער: נאך פארשונג – SwitchHead** – רעדוצירט דעם זכרון און די חשבונות וואס טראנספארמערס דארפן, אן פארלירן קוואליטעט. "Object identifiers" – העלפן שפראך-מאדעלן פארשטיין 3D סצענעס. LLF-LUT – פארבעסערט HDR בילדער.

• **18 דעצעמבער: Meta's ווידעא דאטא-סעט** – Meta האט ארויסגעלאזט א ריזיגן דאטא-סעט פון ווידעאס געכאפט פון דעם באניצער'ס אייגענע אויגן. ער אנטהאלט דאטא פון אסאך סענסארן און האנט-פאזיציעס, און גיט באזיס-רעזולטאטן פאר עקזיסטירנדע מעטאדן. לויט דער ידיעה זענען קוואליטעטישע דאטא-סעטס די הויפט סיבה פארוואס ווידעא-פארשטאנד הייבט יעצט אן צו געלינגען.

• **18 דעצעמבער: סטאביל טריינירן טראנספארמערס** – טראנספארמערס זענען אפט שווער צו טריינירן צוליב אומסטאביליטעט. א הויפט פראבלעם איז ווען די "ענטראפיע" פון אן Attention מאטריקס פאלט צוזאמען. א נייע ארבעט שלאגט פאר א פשוט'ע איבער-פאראמעטריזאציע כדי דאס צו פארמיידן.

- **18 דעצעמבער: DriveMLM פאר זעלבסט-פארנדיגע קארס** – DriveMLM ניצט גרויסע שפראך-מאדעלן פאר אויטאנאמע פארן. ער פארבינדט באשלוסן אין שפראך מיט קאנטראל פונעם קאר, לאזט זיך איינבויען אין עקזיסטירנדע סיסטעמען, און ווייזט בעסערע רעזולטאטן אין סימולאציעס.
- **18 דעצעמבער: DiffusionLight שאצט ליכט** – א נייע טעכניק שאצט אפ די באלייכטונג אין בילדער אן HDR פאנאראמא דאטא-סעטס. זי ניצט דיפוזיע-מאדעלן כדי אריינצומאלן כראם-קוילן אין געווענליכע בילדער.
- **19 דעצעמבער: גראף-נעצן און סעגמענטאציע** – DSF פארבעסערט ספעקטראלע גראף נירא-נעצן מיט פילטער-וואגן ספעציפיש פאר יעדן קנופ, וואס העלפט ביי קאמפליצירטע נעצן ווי דאס וועב. SmooSeg איז א מעטאד פאר סעמאנטישע סעגמענטאציע, באזירט אויפן פרינציפ אז שכנישע טיילן אין א בילד האבן געווענליך ענליכע באדייטונג.
- **19 דעצעמבער: Anthropic: דיסקרימינאציע אין מאדעלן** – Anthropic באשרייבט א מעטאד אפצושאצן פון פאראויס צי א שפראך-מאדעל קען זיין דיסקרימינירנד. מען שאפט א ברייטע רייע פראמפטס פאר באשלוס-סצענארן, מיט פארשידענע דעמאגראפישע פרטים. די וועגן צו פארמינערן דיסקרימינאציע, סיי פאזיטיווע סיי נעגאטיווע, זענען בעיקר פארויכטיגע פראמפט-אינושענירע.
- **19 דעצעמבער: שנעלערע ארכיטעקטורן** – א פארשונג האט געפונען אז UNet ענקאדערס אין דיפוזיע-מאדעלן ביהייבן זיך סטאביל אין פארגלייך מיט די דעקאדערס. דאס האט געפירט צו א שיטה וואס פארשנעלערט באדייטנד טעקסט-צו-בילד און טעקסט-צו-וידעא. די שאפער פון Mamba האבן באשריבן פשוט'ע "sequence mixing" ארכיטעקטורן וואס זענען אסאך שנעלער ווי נארמאלע טראנספארמערס.
- **20 דעצעמבער: וויזשן פארשונגען** – א מולטי-מאדאלער מאדעל פאר נאכגיין אביעקטן ניצט די מעלות פון סיי RGB סיי אינפרא-רויט. א פארשונג ניצט פיר פאר-טריינירטע CNN מאדעלן כדי צו דערקענען פאלשע (deepfake) פנימער אין וידעאס מיט הויכער פינקטלעכקייט. STSS פארבעסערט באדייטנד ריל-טיים רענדערינג מיט הויכער רעזאלוציע און אסאך בילדער א סעקונדע.
- **20 דעצעמבער: M3DBench און Tokenize Anything** – M3DBench איז א דאטא-סעט מיט איבער 320,000 אנווייזונג-און-ענטפער פארן, וואס פאראייניגן טעקסט, בילדער און 3D אביעקטן. Tokenize Anything איז א מאדעל וואס טוט צוגלייך סעגמענטאציע, דערקענונג און באשרייבונג פון בילדער.
- **21 דעצעמבער: די Gemini פארשונג-ארבעט** – Google האט פובליקירט א פארשונג-ארבעט מיט איבער 850 מחברים איבער איר Gemini ארבעט. זי גיט ווייניג פרטים איבער דער ארכיטעקטור און דעם דאטא-סעט, אבער שטעלט צו קרעדיט און רעזולטאטן פון אפשאצונגען.
- **21 דעצעמבער: 3D, שריפטן און טעקסט אין בילדער** – 3D-LFM ניצט טראנספארמערס צו רעקאנסטרווירן 3D סטרוקטורן פון 2D פונקטן, אן פאסיגע 3D דאטא. לויט דער ידיעה איז דאס דער ערשטער פון זיין סארט. FontDiffuser's Alibaba שאפט שריפטן פון איין ביישפיל און איז בפרט גוט מיט קאמפליצירטע אותיות. Diff-Text שאפט רעאליסטישע בילדער מיט טעקסט אין יעדער שפראך, אן טריינירן.
- **21 דעצעמבער: AI גובר אויף Labyrinth** – אן AI שפילט דאס פיזישע שפיל Labyrinth פיל בעסער ווי מענטשן, מיט א בעסטע צייט פון 14.8 סעקונדעס. ער האט זיך עס אויסגעלערנט אין בלויז זעקס שעה.

- **22 דעצעמבער: Apple'ס פארשונג** – Apple האט פובליקירט צוויי פארשונג-ארבעטן. איינע באשרייבט א נייע טעכניק פאר 3D אוואטארן, די אנדערע א מער עפעקטיווע וועג צו לויפן שפראך-מאדעלן. די טעכניק קענען דערמעגלעכן AI סיסטעמען צו לויפן אויף iPhones און iPads.
- **22 דעצעמבער: PowerInfer אויף היימישע GPUs** – לויט דער פארשונג-ארבעט האבן מאדעלן "הייסע" נוראנען וואס ווערן גענוצט כמעט פאר יעדן אינפוט, און "קאלטע" וואס ווערן גענוצט אסאך זעלטענער. אז מען לאדנט פאראויס די הייסע נוראנען אויף דער GPU, שפארט מען זכרון אן צו פארלירן פיל שנעלקייט. עס איז דא א באגלייטנדע קאוד-ביבליאטעק.
- **22 דעצעמבער: בילד-פארשונגען** – א "sequence-aware" פארלוסט-פונקציע פארבעסערט די בילד-קוואליטעט פון דיפוזיע-מאדעלן אויף CIFAR10 און CelebA. "Splatter Image" רעקאנסטרווירט 3D פון איין איינציגן בילד, אנשטאט די איבער 200 בילדער וואס Gaussian Splatting דארף אפט. CLIP-DINOiser טוט סעמאנטישע סעגמענטאציע אן באצייכנטע דאטא. UDiffText פארמינערט אויסלייג-טעותים אין טעקסט אין AI בילדער.

געצייג פאר פראגראמירער און אפן-סאורס

- **4 דעצעמבער: aiPDF** – א געצייג צו קיצן, אנאליזירן, און "רעדן" מיט סיי וועלכע טעקע.
- **5 דעצעמבער: עטליכע נייע געצייג** – א בלאג פונעם שאפער פון React Native ווייזט ווי צו באהאנדלען שפראך-מאדעלן אין קאוד ווי פונקציעס מיט א קלארן "אריין" און "ארויס". AIConfig פון LastMile (אפן-סאורס) – לאזט לייכט טוישן צווישן מאדעל-פארזארגער. Microsoft – א קורס פון 12 לעקציעס איבער גענעראטיווע AI פאר אנהייבער. Unsloth – שנעלער און באלעבאטישער מיט זכרון ביים טריינירן מאדעלן. SchoolXpress – מאכט ביכער און נאטיצן פאר אינטעראקטיווע לערן-מאטעריאל.
- **6 דעצעמבער: Vector Art** – שאפט "וועקטאר" בילדער (בילדער וואס בלייבן שארף אין יעדע גרייס) מיט AI.
- **7 דעצעמבער: Apple'ס MLX** – Apple האט שטילערהייט ארויסגעלאזט MLX, א פשוטע סיסטעם פאר AI חשבונות. זי לויפט שנעל אויף Mac'ס צוליב זייער געמיינזאמען זכרון, און ניצט אויך דעם GPU. א טאג שפעטער האט מען שוין געלאפן Llama 2 אויף MLX, און עס איז ארויס א ביישפיל פון Mistral אויף MLX.
- **7 דעצעמבער: עטליכע נייע געצייג** – Stable Fast – א שנעלע מאשין צו לויפן דיפוזיע מאדעלן פון Hugging Face. OpenAI Switch Kit – לאזט איבערגיין צו אפן-סאורס מאדעלן מיט עטליכע שורות קאוד. CopilotKit – בויט טשעטבאטס אין React אפליקאציעס. EfficientSAM פון Meta – א פיל שנעלערע ווערסיע פון Segment Anything, א מאדעל וואס צעטיילט בילדער.
- **8 דעצעמבער: Mamba Chat און נאך** – Mamba איז א מאדעל וואס איז נישט באזירט אויף טראנספארמערס און ארבעט זייער גוט. מען האט אים צוגעפאסט פאר טשעט, און ער געפינט אינפארמאציע פיל בעסער ווי טראנספארמערס פון דער זעלבער גרייס. SEO GPT – פארבעסערט א וועבזייטל'ס SEO (דאס ווי הויך עס קומט ארויס אין זוך-רעזולטאטן) אינעווייניג אין ChatGPT. UIDraw – צייכנט און בויט א וועבזייטל פונעם טעלעפאן.
- **11 דעצעמבער: עטליכע נייע געצייג** – א טיפע איבערבליק איבער די טעכניקן צו לויפן שפראך-מאדעלן ביז 100 מאל שנעלער. Pearl פון Meta – א ביבליאטעק פאר AI אגענטן מיט reinforcement learning,

- גרייט פאר עכטע ניץ. Ollama – לויפט לאקאלע מאדעלן גאר לייכט, אפילו אויף א CPU. א וויזואליזאציע ווייזט ווי nano-gpt מיט 85 טויזנט פאראמעטערס סארטירט זעקס אותיות. Giskard – טעסטירט מאדעלן.
- **12 דעצעמבער: עטליכע נייע געצייג** – BricksLLM – א "gateway" (א פארמיטלער) פאר OpenAI, געשריבן אין Go, מיט API שליסלען וואס האבן לימיטן אויף צאל רופן, קאסטן און צייט. KwaiAgents – א סיסטעם פאר AI אגענטן. א לעקציע איבער פראמפטס אין DALL-E 3. א רעדע איבער די נייעס אין PyTorch 2. Transformers 4.36.0 פון Hugging Face – מיט שטיצע פאר Mistral און AMD, און safetensors אלס סטאנדארט. Unsloth – שרייבט אליין, אין Triton, די חשבונות פאר טריינירן, מיט א טענה פון 30 מאל שנעלער (די בענטשמארקס זענען נישט די איבערצייגנדסטע). AI Tamago – א לאקאלער מאדעל אלס א "פעט" (א ליבלינג-חיה). Marketing GPTs – א מארקעטינג געהילף.
 - **13 דעצעמבער: עטליכע נייע געצייג** – EasyVolcap – א ביבליאטעק פאר פארשונג אין 3D ווידעא. Respell – אויטאמאטיזירט ארבעט מיט AI אגענטן. Ollama – לייגט צו שטיצע פאר מאדעלן מיט בילדער, ווי Llava.
 - **14 דעצעמבער: Cerebras gigaGPT און נאך** – Cerebras, וואס מאכט טשיפס פאר טריינירן, האט ארויסגעלאזט א פשוטע סיסטעם צו טריינירן ריזיגע מאדעלן, אן צוטיילן זיי צווישן עטליכע GPU'ס. Anyscale האט צוגעלייגט JSON מאד און function calling צו Anyscale Endpoints: דער מאדעל'ס ענטפער קומט אין א פעסטער פארעם, און ער קען רופן API'ס. Swiftly – אן AI געהילף פאר געשעפט-רייזעס.
 - **15 דעצעמבער: לערן-מאטעריאל און געצייג** – א "וועג-קארטע" צו לערנען גענעראטיווע AI אין 2024, פאר אנהייבער. Dover – קוקט איבער אפליקאציעס פאר ארבעט 10 מאל שנעלער. א זאמלונג פון טעגליכע געוואוינהייטן צו פארבעסערן AI ידיעות.
 - **18 דעצעמבער: PyTorch, Coffee און DeepEval** – נייע געצייג אין PyTorch 2.1 ווייזן ווי מען ניצט דעם זכרון, און קענען עס אפילו צעטיילן אין באקאנטע קאטעגאריעס. Coframe'ס Coffee העלפט מיט AI ביי פראנטענד ארבעט; די מאנשאפט שאפט אליין איבער 80% פון איר פראנטענד מיט DeepEval AI. איז א ראם פאר אפשאצן אפליקאציעס מיט שפראך-מאדעלן, מיט איבער 14 מאסשטאבן ווי האלוצינאציע און RAGAS.
 - **18 דעצעמבער: Medprompt+ פראמפטס** – Microsoft'ס Medprompt+ איז דער בעסטער פראמפט-מעטאד פאר דעם MMLU בענטשמארק. Microsoft האט ארויסגעגעבן די מוסטער-פראמפט צוזאמען מיט אנדערע chain-of-thought מוסטערן.
 - **19 דעצעמבער: עטליכע נייע געצייג** – Amphion – אן אפן-סאורס געצייג פאר שאפן מוזיק און רעדע. Google Big Vision – לייגט צו weight sharding פאר טריינירן אויף עטליכע מאשינען. Jaxtyping – כאפט פארם-טעותים אין קאוד דורך זיי צו קאנטראלירן אלס טייפס. Microagents – AI אגענטן וואס פאררעכטן אליין זייערע פראמפטס און Python קאוד.
 - **20 דעצעמבער: Distil-Whisper און LLMingua** – Distil-Whisper איז 6 מאל שנעלער און 49% קלענער ווי Whisper, און בלייבט אין LLMingua. WER. 1% דריקט צוזאם פראמפטס ביז 20 מאל, מיט מינימאלן פארלוסט אין רעזולטאטן.

- **21 דעצעמבער:** **TinyZero, MossFormer2, WhisperPlus** – טריינירט AlphaZero – ענליכע אנגענטן אויף יעדע סביבה. MossFormer2 – צעטיילט בעסער קולות פון איין מיקראפאן. WhisperPlus – פארגעשריטענע רעדע-צו-טעקסט.
- **22 דעצעמבער:** **Llava, Llama Recipes, Whisper Turbo** – Llava ערלויבט יעצט א שמועס וואס מישט טעקסט, בילד-אינפוט און בילד-אויסגאבע. Llama Recipes זענען געצייג און ביישפילן פאר ניצן Meta's Llama מאדעלן. Whisper Turbo איז 6 מאל שנעלער ווי נארמאלער Whisper.
- **18 דעצעמבער:** **Mistral 7B** אויף **Magic** – אן עקספערימענט האט פיינגעטונט א מאדעל מיט 7 ביליאן פאראמעטער אויף "drafts" פונעם קארטן-שפיל Magic: The Gathering. דער מאדעל איז געקומען נאנט צו מענטשלעכער שטאפל.

פירמעס, געלט און ביזנעס

- **4 דעצעמבער:** **Altman צוריק ביי OpenAI** – אין אן אינטערוויו האט Altman נישט גערעדט איבער די סיבות פארוואס מען האט אים ארויסגעווארפן. ער האט דערציילט אז ער האט זיך ערשט געשפירט טראצעדיג, און שפעטער אנגענומען צוריקצוקומען. Microsoft וועט נעמען א זיץ אויפן באורד אן א שטים. OpenAI's GPT Store איז אפגעשטופט געווארן ביז נעקסטן יאר.
- **4 דעצעמבער:** **Stability AI אין שוועריקייטן** – Stability AI האט געהאלטן ערשטע שמועסן מיט עטליכע פירמעס וועגן פארקויפן זיך, צוליב דרוק פון אינוועסטירער איבער איר פינאנציעלן מצב. אין Sifted שרייבט מען אז די פירמע כאפט זיך צו יעדן שטרוי, ווען זי טוישט אירע פלענער און פירט איין אן אבאנעמענט-פרייז.
- **4 דעצעמבער:** **Gemini אפגעשטופט** – לויט א ערשטן באריכט האט Google אפגעשטופט Gemini ביז יאנואר, ווייל ער האט זיך נישט קאנסיסטענט באהאנדלט מיט פראגעס אין שפראכן אחוץ ענגליש.
- **5 דעצעמבער:** **OpenAI און Rain AI טשיפס** – OpenAI האט מסכים געווען צו קויפן טשיפס פאר 51 מיליאן דאלאר פון Rain AI, א סטארטאפ אין וועלכן Altman האט פערזענליך אינוועסטירט. דאס ווייזט ווי Altman's ראלעס ווי אינוועסטירער און ווי CEO זענען פארפלאכטן, א ענין וואס האט צוגעהאלפן צו זיין אפזאגן.
- **6 דעצעמבער:** **xAI זוכט 1 ביליאן** – Musk's xAI האט זיך איינגעגעבן ביי דער אמעריקאנער אויפזיכט פאר "securities" (פינאנציעלע פאפירן) כדי אויפצונעמען ביז 1 ביליאן דאלאר.
- **6 דעצעמבער:** **קויפן און געלט-רונדעס** – Airbnb האט געקויפט GamePlanner.AI, וואס ווערט געפירט דורך Adam Cheyer (מיט-גרינדער פון Siri) און Siamak Hodjat. Solve Intelligence, וואס העלפט אדוואקאטן שרייבן פאטענטן, ווערט שוין גענוצט ביי איבער 25 פירמעס, מיט באריכטן פון 60 ביז 90% מער עפעקטיווקייט. זי האט אויפגענומען 3 מיליאן דאלאר.
- **6 דעצעמבער:** **פארוואס מען האט אפגעזאגט Altman** – לויט א באריכט האט מען Altman באשולדיגט אז ער האט מאניפולירט באורד-מיטגלידער כדי ארויסצושטופן א מיטגליד וואס האט קריטיקירט OpenAI's צוגאנג צו זיכערהייט. א צווייטער באריכט זאגט אז הויכע פירער ביי OpenAI האבן געווארנט דעם באורד איבער Altman's אויפפירונג, אז ער האט געשאפן כאאס און אויפגעהעצט ארבעטער איינער קעגן צווייטן, און איבער אומערליכקייט קעגן'ם באורד. ער איז צוריקגעקומען דאנק די שטארקע שטיצע פון די ארבעטער, וואס האבן געסטראשעט מיט מאסן-רעזיגנאציעס.

- **7 דעצעמבער: Musk האט געדריקט OpenAI** – באלד איידער Musk האט פארלאזט OpenAI, האט ער זיי געבעטן ארבעטן שנעלער. דאס איז אנדערש פון זיינע געווענליך מער סקעפטישע רייד איבער AI.
- **8 דעצעמבער: Google TPU v5p** – Google האט אנאנסירט TPU v5p, איר שטארקסטע טשיפ פאר AI ביז היינט. אויך אנאנסירט: AI Hypercomputer, א סופער-קאמפיוטער ארכיטעקטור מיט ייזנווארג, אפענע סאפטוועיר, און בייגעוודיגע מעגליכקייטן ווי צו באצאלן.
- **8 דעצעמבער: Brad Lightcap רעדט** – OpenAI'ס COO רעדט איבער ChatGPT, איבער Dev Day, און איבער ווי Altman טראכט.
- **11 דעצעמבער: Liquid AI** – א נייער סטארטאפ פון MIT בויט "liquid neural networks". זיי זענען פיל קלענער ווי געווענליכע מאדעלן און דארפן פיל ווייניגער קאמפיוטער-כח.
- **12 דעצעמבער: Microsoft און די יוניאנס** – Microsoft ארבעט צוזאמען מיט דער AFL-CIO איבער AI'ס איינפלוס אויף ארבעטער, און וועט געבן AI טריינינג פון ווינטער 2024. דער אפמאך שליסט איין נייטראליטעט, וואס מאכט גרינגער פאר ארבעטער ביי Microsoft זיך צו פאראייניגן אין א יוניאן.
- **14 דעצעמבער: a16z גרענטס** – די צווייטע רונדע פון a16z'ס גרענטס פאר אפן-סאורס AI איז געגאנגען צו פראיעקטן ווי Axolotl און Lmsys.
- **14 דעצעמבער: OpenAI און Axel Springer** – OpenAI וועט טריינירן אירע מאדעלן אויף אינהאלט פון Axel Springer (Business Insider, Politico), און צולייגן פרישע ארטיקלען צו ChatGPT.
- **15 דעצעמבער: Answer AI** – דער גרינדער פון Fast AI עפנט א לאבאראטאריע פאר פראפיט. זי וועט נישט בויען נייע מאדעלן, נאר זוכן פראדוקטיווע וועגן צו ניצן די עקזיסטירנדע.
- **18 דעצעמבער: AMD MI300X** – לויט AMD'ס בענטשמארק איז איר הויפט AI טשיפ MI300X, מיט ROCm 6, יעצט נאנט צו Nvidia ביי אינפערענס. דאס עפנט אן אלטערנאטיווע פאר נייע AI פירמעס.
- **18 דעצעמבער: Agility'ס ראבאט** – Agility'ס צוויי-פייסגער ראבאט Digit קען יעצט אויספירן קאמפליצירטע ארבעטן לויט באפעלן אין נאטירלעכער שפראך, מיט'ן ניצן שפראך-מאדעלן.
- **19 דעצעמבער: ByteDance און OpenAI** – ByteDance, די מוטער-פירמע פון TikTok, האט בסוד גענוצט OpenAI'ס טעכנאלאגיע, אקעגן די באדינגונגען, כדי צו בויען א קאנקורירנדן מאדעל, "Project Seed". אינערליכע קאמוניקאציעס ווייזן אויף פלענער "אפצוואוואשן" די באווייזן.
- **20 דעצעמבער: Harvey: \$80 מיליאן** – דער לעגאלער סטארטאפ Harvey, געבויט אויף OpenAI טעכנאלאגיע, האט אויפגענומען \$80 מיליאן אין א Series B רונדע, ביי א ווערט פון איבער \$700 מיליאן. די פירמע בויט מאדעלן מיט OpenAI פאר געזעץ און אדוואקאטור.
- **20 דעצעמבער: Anthropic: באדינגונגען און געלט** – Anthropic האט דערהיינטיגט אירע באדינגונגען איבער וועם עס געהערט וואס דער מאדעל שאפט. די פירמע זאגט אויך אז זי וועט פארטיידיגן אירע קונים אין געריכט קעגן קאפירעכט קלאגעס. לויט באריכטן פירט Anthropic שמועסן צו אויפנעמען \$750 מיליאן ביי א ווערט פון \$15 ביליאן, אין א רונדע געפירט פון Menlo Ventures.
- **20 דעצעמבער: Stability AI אבאנעמענט** – Stability AI פירט איין א באצאלטע מעמבערשיפ פאר קאמערציעלן באנוץ פון אירע מאדעלן.

• **20 דעצעמבער: פליגער אן פילאט** – Reliable Robotics האט געפלוין א קליינעם פראכט-פליגער פון FedEx אן א מענטש אויפן ברעט, אין א פלי פון 12 מינוט. דאס איז געווען איר צווייטער אויטאמאטישער פלי. דער סטארטאפ ארבעט צו באקומען פולע דערלויבעניש פון FAA. יעדער ווייטער פילאט וועט אויפפאסן אויף נאר איין פליגער אויף אמאל.

• **21 דעצעמבער: Waymo קארס** – Waymo's פארשער זאגן אז אירע קארס אן דרייווער האבן באדייטנד ווייניגער עקסידענטן וואס ווערן געמאלדן צו פאליציי אדער פירן צו וואונדן, אין פארגלייך מיט מענטשלעכע דרייווערס. די פירמע שטופט פאר אן אינטערנאציאנאלן סטאנדארט אין אנאליזירן אזעלכע דאטא.

• **18 דעצעמבער: קליינע AI אפליקאציעס** – PriceGPT – אנאליזירט אייער אדער א קאנקורענט'ס פרייז-בלאט. Facts& – פראדוקט-פארשונג פאר DTC מארקעס. AI – Outboundly פאר "קאלטע" אימעילס. EnVSION AI – פאסט צוזאמען שעהן פון קונה-אינטערוויוס אין מינוטן.

רעגירונגען, געזעצן און זיכערהייט

• **5 דעצעמבער: AI און צוטרוי** – א רעדנער ווארנט אז מענטשן באטראכטן AI ווי פריינט אנשטאט ווי א סערוויס, און פירמעס וואס זוכן פראפיט קענען דאס אויסניצן. ער פאדערט אז די רעגירונג זאל אריינמישן מיט געזעצן פון דורכזיכטיגקייט, און מען זאל אנטוויקלען פובליקע AI מאדעלן.

• **6 דעצעמבער: AI Alliance** – IBM און Meta פירן אן אליאנס פון איבער 50 ארגאניזאציעס פאר אפענע און פאראנטווארטליכע AI: סטאנדארטן, יינווארג, און ערציאונג.

• **6 דעצעמבער: פראטאקאל פאר סטארטאפס** – Responsible Innovation Labs האט ארויסגעלאזט א פרייוויליגן פראטאקאל פאר פאראנטווארטליכע AI ביי סטארטאפס. 35 "venture capital" פאנדס (וואס אינוועסטירן אין סטארטאפס) שטיצן אים, און ער פאסט זיך צו מיט דער Biden אדמיניסטראציע'ס זעאונג.

• **8 דעצעמבער: Purple Llama** – Meta's זיכערהייט-פראיעקט מיט פרייע געצייג: CyberSecEval צו אפשאצן די זיכערהייט פון AI סיסטעמען, און Llama Guard. ער כאפט אויך אומזיכערן קאוד.

• **11 דעצעמבער: FTC קוקט אויף Microsoft** – די FTC (די אמעריקאנער אויפזיכט אויף האנדל) אונטערזוכט Microsoft's באציאונג צו און אינוועסטירונג אין OpenAI.

• **12 דעצעמבער: EU AI Act** – די אייראפעאישע יוניאן האט דערגרייכט א פרעלימינארן אפמאך אויף דעם ערשטן פולשטענדיגן AI געזעץ. ער באהאנדלט גענעראטיווע AI, און די פאליציי'ס ניץ פון געזיכט-דערקענונג. דער געזעץ טיילט איין AI סיסטעמען לויט זייער ריזיקע, מיט צוויי שטאפלן פאר אלגעמיינע מאדעלן און פארבאטן פאר "אומאנעמבארע" ריזיקע. די לאנגע פארהאנדלונגען זענען געווען שווער, ספעציעל איבער געזיכט-דערקענונג און יסוד-מאדעלן. עס קען ווערן א מוסטער פאר AI געזעצן איבער דער גאנצער וועלט.

• **13 דעצעמבער: נאך וועגן אויפזיכט** – Interoperable Authorization Protocol – אן אפן-סאורס פראטאקאל פאר זיכערע קאמוניקאציע צווישן מאדעלן און באניצער, מיט א סיסטעם פאר דערלויבעניש. א קאמיטעט פון MIT שלאגט פאר צו פארברייטערן די עקזיסטירנדע רעגולאציעס און פאראנטווארטליכקייט אויף AI.

- **22 דעצעמבער: Rite Aid פארבאטן** – Rite Aid האט גענוצט פנים-דערקענונג וואס האט פאלש אידענטיפיצירט מענטשן אלס גנבים, און דאס האט געפירט צו דראמאטישע קאנפראנטאציעס. יעצט איז די פירמע פארבאטן צו ניצן די טעכנאלאגיע. די FTC ווערט שטרענגער מיט אזעלכע טעכנאלאגיעס.
- **22 דעצעמבער: OpenAI זיכערהייט** – OpenAI האט געשאפן א נייע זיכערהייט-ראטגעבער גרופע. דער באורד האט באקומען וועטא-מאכט איבער אלע מאדעלן.

אנאליזן און מיינונגען

- **4 דעצעמבער: סטראטעגיע און Copilot – Not Boring** – ברענגט די לעצטע געשעענישן ביי OpenAI אלס א ביישפיל פון א גוטער "grand strategy". GitHub'ס הויפט פראדוקט-מנהל זאגט אז AI וועט נישט פארבייטן פראגראמירער, נאר זיי מאכן עפעקטיווער. Matt Rickard שרייבט אז GitHub "גרינדעט איבער" Copilot, און עס קען אפשר איבערבויען גאנצע ארבעט-פראצעסן.
- **5 דעצעמבער: קריטיק אויף "doomers"** – אן ארטיקל טענהט אז די גרופע ביי OpenAI וואס האט מורא פון א AI קאטאסטראפע ("doomers") האט געמאכט סטראטעגישע טעותים, און דערמיט אדרבה געיאגט די אנטוויקלונג פון AI.
- **6 דעצעמבער: Hoffman און Khosla – Reid Hoffman און Vinod Khosla** – רעדן איבער די צוקונפט פון AI, "טעכנא-אפטימיזם", און ווי אזוי מען זאל רעגולירן די אינדוסטריע.
- **7 דעצעמבער: אפטימירן מאדעלן און אויגמענטאציע – Lightspeed** – טיילט עצות פון Google און TitanML צו פארבעסערן מאדעלן, סיי דורך פראמפטס און סיי דורך נאכטריינירן. א אנדער ארטיקל ראט צו טראכטן וואס צו אויטאמאטיזירן אין דער גאנצער ראלע, נישט נאר זיך העלפן טון די אייגענע ארבעט.
- **11 דעצעמבער: Gemini'ס מארקעטינג** – דער מחבר פון The Intrinsic Perspective שרייבט אז די אקציעס פון Google זענען קוים ארויף נאך דער דעמא, און אז די אינדוסטריעס וואס AI טוישט זענען נישט אזוי פראפייטאבל. Big Technology פרעגט פארוואס Google האט געדארפט "אויפבלאזן" איר מארקעטינג, און זאגט אז דאס ווייזט ווי פיל עס שטייט אויפן שפיל אין א געיעג וואו דער געווינער נעמט אלעס.
- **13 דעצעמבער: קיין נייער פארשפרייטונג-וועג** – לויט א שמועס ביים TCV Engage Summit פעלט AI א נייער וועג צו דערגרייכן קונים. דאס גיט א מעלה צו עטאבלירטע פירמעס.
- **14 דעצעמבער: "Vertical AI"** – א שותף ביי Greylock זאגט פאר אז באניצער וועלן ערווארטן סאפטוועיר געמאכט ספעציעל פאר זייער פאך ("vertical"), נישט אלגעמיינע ("horizontal") סאפטוועיר.
- **15 דעצעמבער: Q* שמועות** – לויט Understanding AI איז Q* (א געשמועסטער געהיימער OpenAI פראיעקט) קוים דער גרויסער דורכבראך צו AGI (AI אויפן שטאפל פון א מענטש). אבער עס קען זיין א שריט צו AI מיט אלגעמיינע טראכט-פעאיגקייטן.
- **18 דעצעמבער: וואו שטייט AI** – פירנדע גרינדער און אינוועסטירער רעדן איבער וואס עס טוט זיך אין AI, פון שפראך-מאדעלן ביז חידושים אין גרויסע פירמעס, אלס א קורצע איבערזיכט.

- **19 דעצעמבער:** געזעץ-טעק און Nvidia'ס קאנקורענץ – Lightspeed גיט איר בליק אויף AI אין לעגאלער טעכנאלאגיע. אן אנאליז זאגט אז מיר זענען קוים א יאר אין דער נייער וועלט, אז די אינטערעסאנטסטע טשיפס זענען נאך נישט אנגעקומען, און אז אלע צילן אויף Nvidia.
- **21 דעצעמבער:** Meta'ס CTO – Andrew Bosworth הויפט-טעכנאלאג רעדט ווי די AI בהלה האט געצוואונגען די פירמע זיך צו ענדערן, און וואהין די טעכנאלאגיע וועט זי פירן.
- **22 דעצעמבער:** מארקעטפלעצער און יאר-סך-הכל – a16z שאצט אז AI וועט מאכן מארקעטפלעצער מער פערזענליך פאר ביידע זייטן. Foundation Capital ציילט אויס זעקס הויפט-טרענדס פון AI אין 2023.

ווייטערע ידיעות

- **8 דעצעמבער:** 300 ביישפילן – א דאטאבאזע מיט 300 ביישפילן פון איבער 80 פירמעס, וואס באשרייבן ווי זיי האבן געבויט סיסטעמען מיט "machine learning" (מאשין-לערנען).
- **13 דעצעמבער:** Project Ellmann – Google באטראכט א פראיעקט וואס זאל ניצן Gemini צו אנאליזירן בילדער און זוכן אויפן טעלעפאן, און דערציילן פאר א מענטש זיין לעבנס-געשיכטע. מען ווייט נאך נישט צי עס וועט אריינקומען אין Google Photos.
- **21 דעצעמבער:** אינזשעניר-געהאלטן – לויט levels.fyi צאלט Jane Street דאס מערסטע פאר אנהייבנדע אינזשענירן: א געהאלט-פעקל פון \$350,000 פאר L1.