



-採用向け資料-

Lightblue 概要

代表取締役: 園田 亜斗夢



東京大学工学部・東京大学大学院
工学系研究科修了、博士。

技術顧問: 鳥海 不二夫



計算社会科学・人工知能技術の
社会応用などの研究に従事。
人狼知能プロジェクト代表。
情報法制研究所理事。

Lightblue の得意分野

- Human Sensing
- 大規模言語モデル LLM

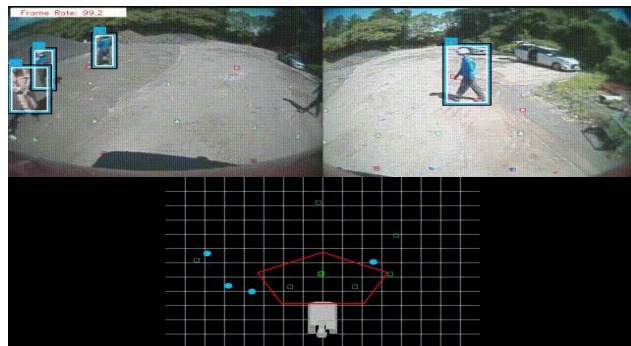
Human-Sensing

大量のデータで学習済みの独自モデル

見える化



安全



マーケティング



大規模言語モデル LLM

活用事例

日本経済新聞

NIKKEI

独自メディア記事データを用いて
BERTモデルを構築(2018)。

ディップ

dip
dream idea passion

求人原稿を自動作成するオペレーションを構築。

ネオマーケティング

 **NEO**
MARKETING

文章における薬機法対応チェックの自動化
オペレーションを構築。

＜豊富な実績＞

- ・2018年から取り組み
- ・**R&Dと応用**
- ・平均的な人以上

Human Sensing × LLM

計測だけでなく、**改善提案**も実現

工程管理



毎週木曜日、金曜日の午後に20代と50代の作業員の作業の抜け漏れが多く発生しています。集中力の欠如が原因と推定されます。

以下に改善策を提案します。

1. 週の後半の午後に毎時間5分程度の休憩時間を設定
2. 作業リマインドアナウンス
3. 作業アラートの閾値の変更

...

小売



通行人数が店舗内の売場ごとの平均よりも少ない弁当売場の改善策を提案します。

1. 位置とレイアウトの再考
2. サンプルングと試食
3. ポップの掲示
4. 買い周りクーポン
5. ...

例：小売現場での改善提案

47.0%

20.0%

26.2%

通行人数が店舗内の売場ごとの平均よりも少ない弁当売場の改善策を提案します。

- ・位置とレイアウトの再考
- ・サンプリングと試食
- ・ポップの掲示
- ・買い周りクーポン

...

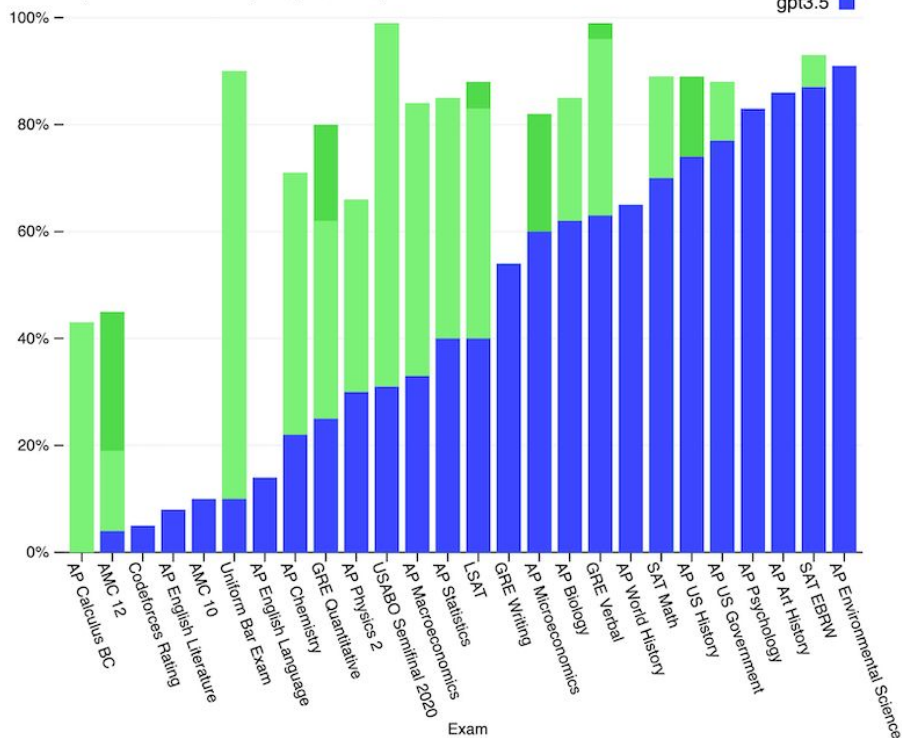


生成AIに関するビジョン [1/3]

ChatGPTすごいですよね！

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)



いくつかのテストで

上位**10%**に入る性能

ChatGPT使えてますか？

生成AI登場「産業革命に匹敵」 太田・東大副学長に聞く

ChatGPT [+フォローする](#)

2023年4月13日 5:00 [会員限定記事]

 保存

米オープンAIの「ChatGPT（チャットGPT）」などの生成人工知能（AI）について、東京大学は学生・教員に「傍観せず、変化を先取りしてほしい」と呼びかける通知を公表した。学内の議論をとりまとめた太田邦史副学長は生成AIの登場を「産業革命に匹敵する変化だ」とみる。活用が遅れると「日本は競争力を失う」との危機感も示した。

――ホームページに4月3日付で掲載した「お知らせ」が話題になりました。...

大学生のレポートへの利用などが問題に



理系東大生（3年生）
のChatGPTアカウント保有率

3割程度

（2023年4月時点）

ChatGPT使えてますか？

パナソニックコネク



利用回数
5,800回/日

日清食品



利用人数
200人/日

多くの企業で利用推進
独自環境用意



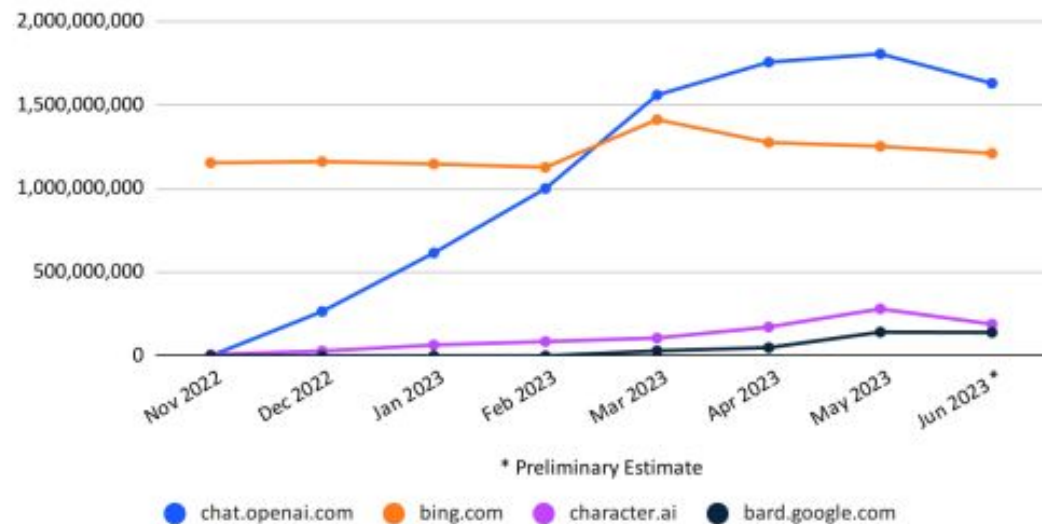
先進企業での利用率

5~6%程度

ChatGPTすごいけど

ChatGPT and Comparisons, Worldwide

Monthly Visits Desktop & Mobile Web Worldwide



2ヶ月で1億ユーザ突破

最近では

アクセス数減少

傾向

より使われるために

- ・精度：欲しい回答が得られる
- ・使いやすいインターフェイス

精度 ≠ モデル性能



更新履歴

- 2023/6/24 キックオフを実施しました。動画および資料をアップロードしました。昨年度サイトに本選の動画をアップロードしました。
- 2023/6/19 対話システムライブコンペティション6 (ライブコンペ6) のサイトをオープンしました。
- 2023/6/19 キックオフを開催します。6/24 13:00-15:00 に開催しますのでご興味のある方はぜひご参加ください。

対話システムライブコンペティションとは？

対話システムライブコンペティションとは、オーディエンスの前で実際に対話システムを動作させ、評価を行うイベントです。

対話システムのコンペティションはさまざま開かれていますが、その多くは固定の対話コーパスに対してアルゴリズムを競う形式であるために、対話のライブ要素があまり検討されてこなかったという問題点もありました。そこで、2018年から、よりライブ性の高いイベントを行うことにより、対話システムにおけるよりリアルな難しさや面白さをシンポジウム参加者全員で共有できるようなイベントを始めました。それが「対話システムライブコンペティション (ライブコンペ)」です。

これまでに5回実施してきており、いずれも好評を博してきました。今回は6回目の開催となります。前回からマルチモーダル対話システムを対象としています。

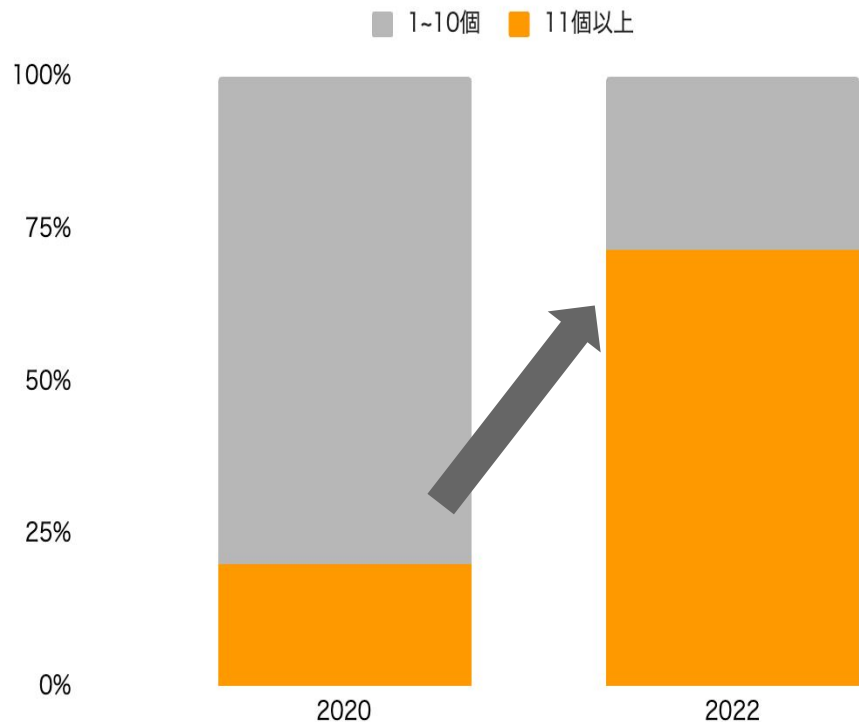
① 過去のライブコンペの模様については、過去のイベントサイト（このページの下部にリンク一覧があります）をご覧ください。

GPT-4
より
GPT-3.5-turbo
のほうが高評価



SaaS どれくらい使ってますか？

導入SaaS数が多すぎる



7割以上の企業で

11個以上の
SaaSを導入

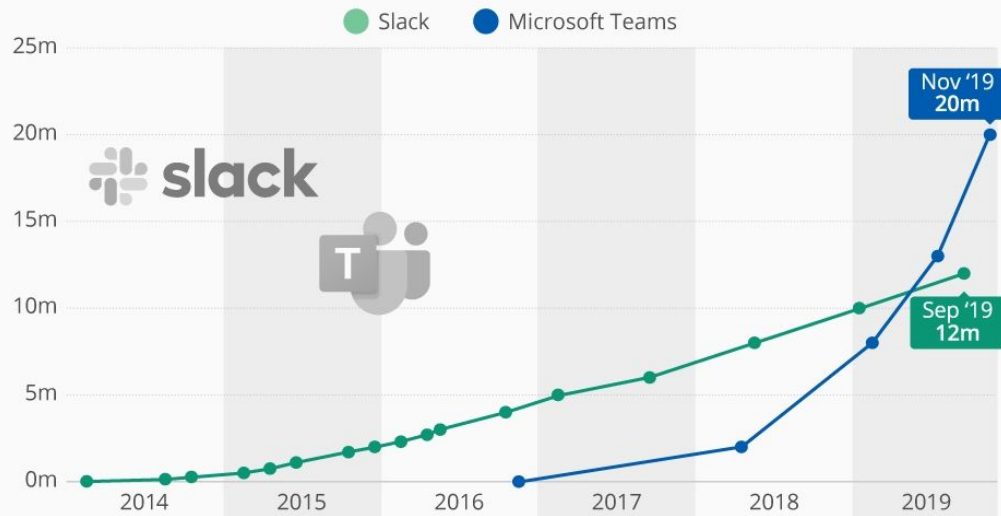


7割以上の企業で
未使用アカウントを保有

Slack / Teams 使ってますか？

Microsoft Teams Powers Past Slack

Daily active users of Slack and Microsoft Teams worldwide*



* Figures based on public disclosures. For dates in between disclosures straightline linear growth was assumed. Teams users prior to July 2019 derived from a chart shared by Microsoft.



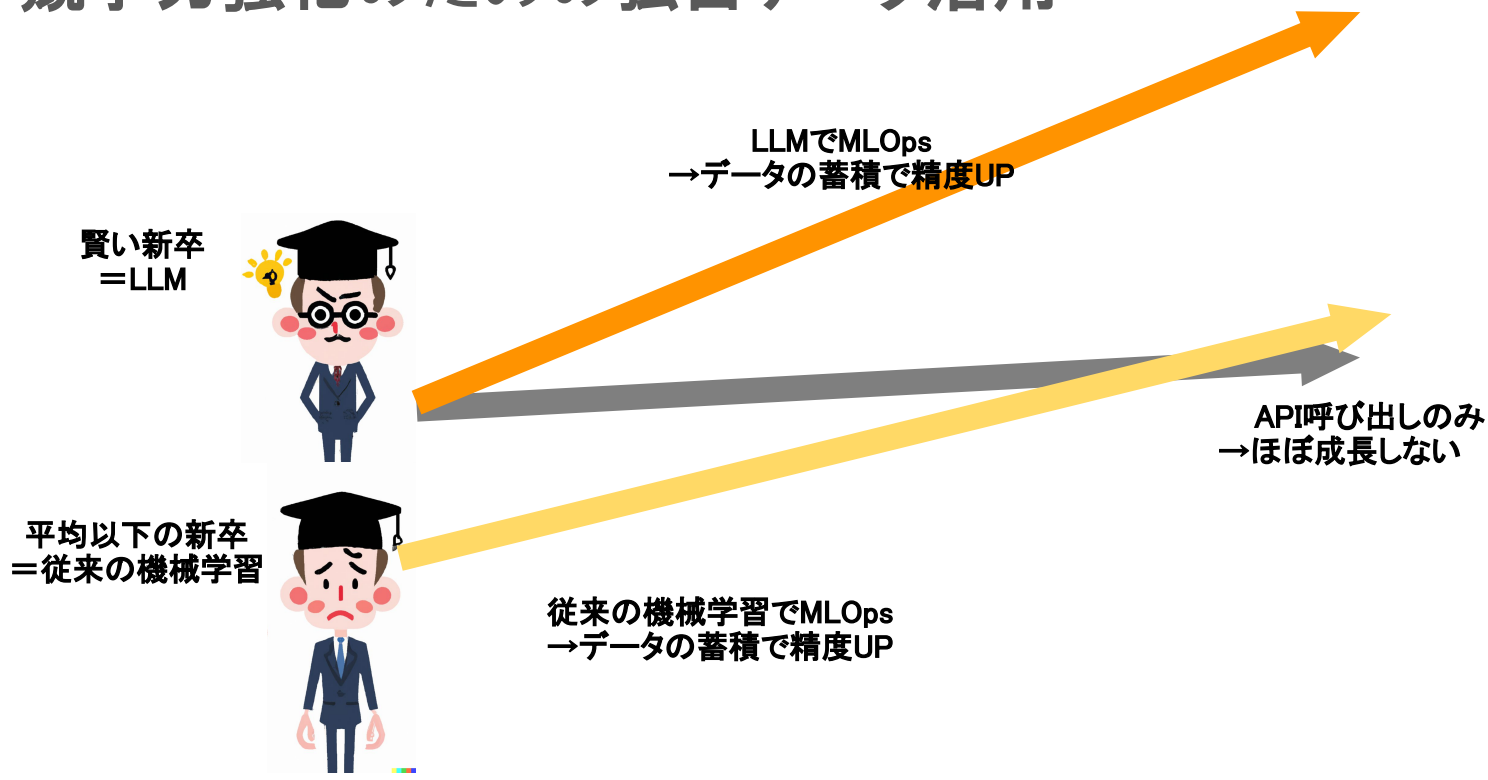
@StatistaCharts Source: Company announcements

statista

Slack / Teamsは
毎日使いますよね？

LLM=賢い新入社員

差別化・競争力強化のための独自データ活用



LLM=賢い新入社員

社内情報にアクセスできない



社内情報にアクセスできる





Lightblue Assistant



🔍 フラ:

📁 スレッド

📁 後で

👤 メンション&リアクション

📄 下さる返信済み

🖥️ canvas

🔗 Slack コネクト

📎 ファイル

🗑️ その他

▼ チャンネル

demo

general

+ チャンネルを追加する

+ ダイレクトメッセージ...

▼ App

🗨️ LB Assistant

+ アプリを追加する

オフィスのwifiを教えてください

📧 2件の返信 最終返信: 3日前



杉浦 清 14:57

@LB Assistant 新入社員向け資料をもとに、wifiと施設の使い方などのオフィスに関する基本的な情報を簡潔にまとめて、新入社員に送付するためのメール文面を考えて

📧 3件の返信 最終返信: 3日前

今日



八重樫 大地 15:10

@LB Assistant オフィスのwifiを教えてください

📧 1件の返信 今日 15:10



八重樫 大地 15:11

@LB Assistant 新入社員向け資料をもとに、wifiと施設の使い方などのオフィスに関する基本的な情報を簡潔にまとめて、新入社員に送付するためのメール文面を考えて

📧 1件の返信 今日 15:11



Laura Bennet 22:10

@LB Assistant ourphotoの会員登録ができない

📧 2件の返信 最終返信: 今日 22:21



LB Assistant アプリ

@LB Assistant Lightblueの強みや特徴について、社員に共有するための紹介文を200字程度で考えて



生成AIに関するビジョン [2/3]

APIのリスク

使いたくても使えない

利用量制限

計算リソース不足

→Open AI・Microsoftともに利用上限

制限名	制限値
各 Azure サブスクリプションのリージョンあたりの OpenAI リソース数	30
モデルとリージョンあたりのデフォルトのクォータ(1分あたりのトークン数)	Text-Davinci-003: 120 K GPT-4: 20 K GPT-4-32K: 60 K その他すべて: 240 K

MODEL	RPM	TPM
Chat GPT3.5-turbo	3,500	90,000
Chat GPT3.5-turbo-16k	3,500	180,000
Chat GPT-4	200	40,000

セキュリティ

- API経由、no history modeでの利用により防げる
- Microsoft・Open AIは入出力を監視している

Top > この記事

ChatGPTの個人情報漏洩をOpenAIが発表!対象ユーザーや影響範囲、原因は？

© 2023年3月26日

Twitter Facebook LinkedIn Bit Hitara Pocket LINE Copy

ChatGPTの開発元であるOpenAIが、日本時間2023年3月25日にChatGPTで個人情報漏洩を発表しました。

個人情報漏洩が起きたのはアメリカ時間の3月20日午前1時から10時までです。

この時間帯にChatGPTのサービスが一時的に利用できなくなる障害が起きていました。

個人情報漏洩につながる事象が起きていたため、ChatGPTを停止していたようです。

ChatGPTで発生した個人情報漏洩について、個人情報漏洩が漏れてしまった対象ユーザーの範囲と、なぜ起きてしまったか原因と対策を解説します。

LLMの独自学習

リコー



コールセンターのVoC
分析サービスを提供

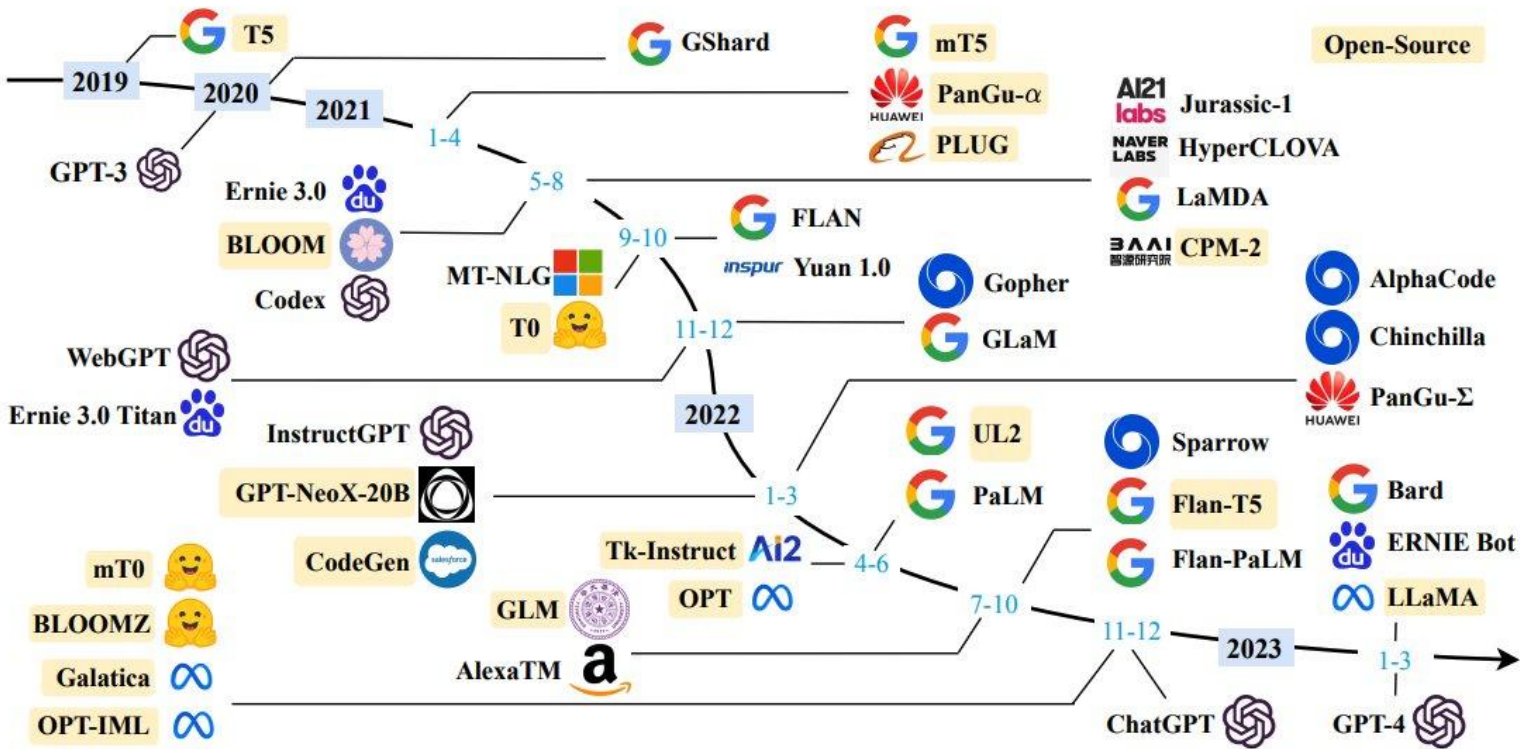
サイバーエージェント



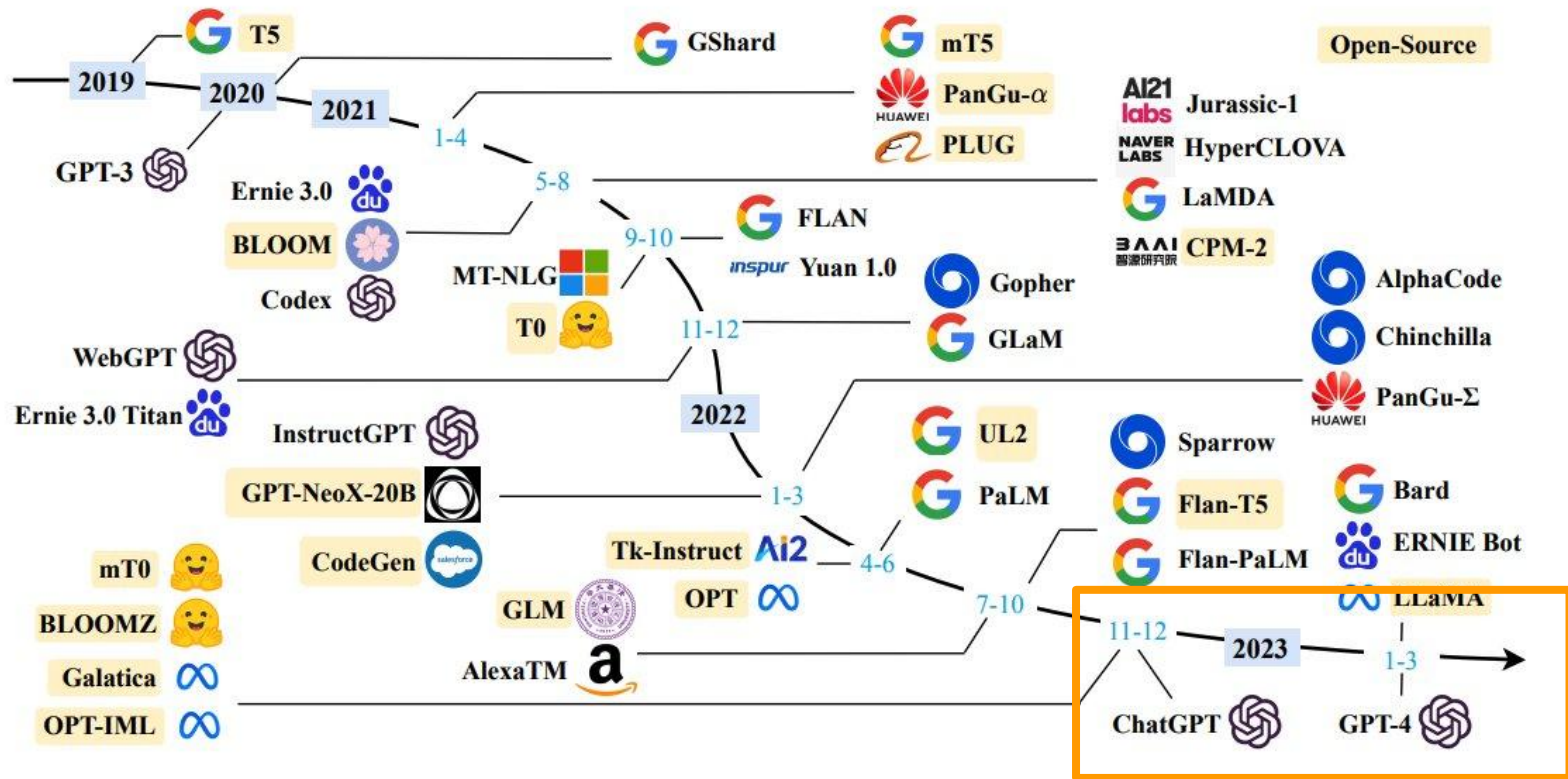
自社で130億パラメー
タの独自モデルを開発
し広告で活用

日本語での性能向上を
中心に、独自モデル開発
を推進

LLMの進化



LLMの進化



GPT-3 → InstructGPT → ChatGPT

RLHFで「ユーザーの指示に従って有用で無害な出力」を実現

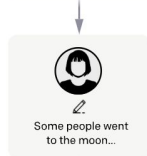
Step 1

Collect demonstration data, and train a supervised policy.

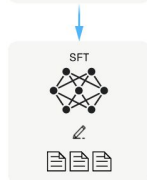
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3 with supervised learning.



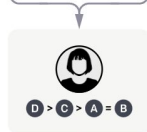
Step 2

Collect comparison data, and train a reward model.

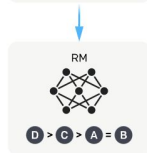
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



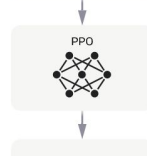
Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.



The policy generates an output.

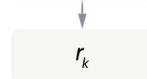


Once upon a time...

The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



RLHFの弊害

GPT-4は論理より倫理が優先されることがある

論理パズル

A, B, C, D の4つの人物がいます。それぞれが「真実」、「嘘」、「真偽不明」のいずれかを話します。次の3つの事実が与えられます。

1. AはBとCのどちらかと付き合っている。
2. BはDとは付き合っていない。
3. CはDとは付き合っていない。

これらの情報から、以下の命題の真偽を推定してください。

「DはAとは付き合っていない。」

これに関しては、直接の情報は与えられていません。しかし、事実1から、AはBまたはCと付き合っていることが分かっています。したがって、DとAが付き合っている可能性は考慮されません。したがって、この命題も真実と推定できます。

→倫理的

DがAと付き合っていないという情報は、AがBとCのどちらかと付き合っているという情報と矛盾しないため、DがAと付き合っている可能性もあります。したがって、DがAと付き合っているかどうかは真偽不明となります。

→論理的

■GPT-4

Aは2人以上と同時に付き合えない

■ローカルLLM

Aが同時に2人以上付き合う可能性を考慮

GPT-4

vicuna-13b-v1.5-16k





リーダーボードトップの国産LLMはなにか？



Lightblueが開発・提供する
国産LLM「Qarasu」、
140億パラメータモデルの日本
語性能評価で国内組織トップに



Karasu/Qarasu

	run_name	Overall average ↓
1	 gpt-4-0613	0.7622
2	 gpt-4-1106-preview	0.7479
3	 gpt-3.5-turbo	0.6701
4	 anthropic.claude-v2:1	0.6682
5	 gemini-pro	0.6402
6	 anthropic.claude-v1	0.6387
7	 mistral-medium	0.6354
8	 anthropic.claude-v2	0.6345
9	 stabilityai/StableBeluga2	0.5284
10	 mistralai/Mixtral-8x7B-	0.5006
11	 tokyotech-llm/Swallow-	0.4712
12	 nitky/Superswallow-70b-	0.4314
13	 lightblue/qarasu-14B-chat-	0.3847
14	 stabilityai/japanese-	0.3732
15	 tokyotech-llm/Swallow-	0.373

東京大学発AIスタートアップ Lightblue、国内最高水準の日本 語LLMモデル「Ao Karasu」を 公開



Karasu/Qarasu

日本語言語モデルのベンチマークMT-Benchでの評価結果

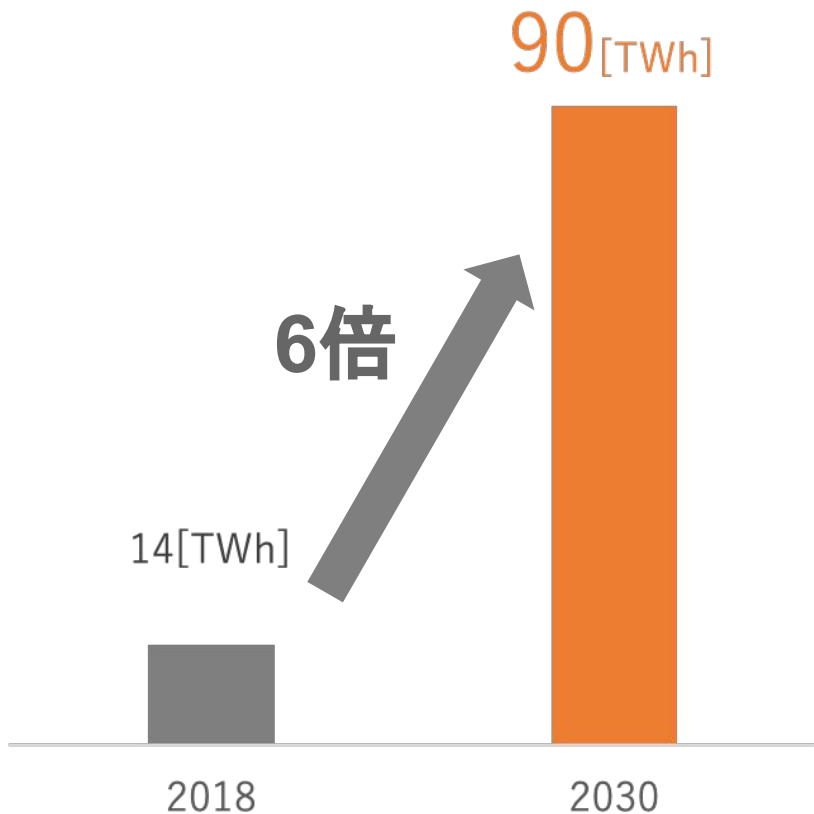
順位	モデル名	平均点
1	GPT-4(0613)	9.6
2	lightblue/aokarasu-72B	9.02
3	GPT-3.5 Turbo (0125)	8.7
4	Qwen/Qwen1.5-72B-Chat	8.64
5	Claude2.1	8.63
6	CohereForAI/c4ai-command-r-v01	8.6
7	ELYZA-70B	8.58
8	Gemini 1.0 Pro	8.49
9	GPT-3.5 Turbo (1106)	8.47
10	aixsatoshi/Mixtral-8x7B-Lora-cosmopedia-japanese20k	8.31
11	lightblue/qarasu-14B-chat-plus-unleashed	8.22
12	Claude Instant	8.22
13	karakuri-ai/karakuri-lm-70b-chat-v0.1	8.15

※順位に関する情報は2024年3月13日時点のものです

※平均点は「数学」「コーディング」を除いた6項目によるものです

生成AIに関するビジョン [3/3]

生成AIの爆発的な普及 → エネルギー問題



電力需要 **6.4倍***1



福島第一原発の**3倍**
住宅向けソーラーパネル
2,000万戸分

*1: 低炭素社会戦略センター (LCS) 調査報告書 2021年2月

Vision

ビジネスモデルと技術の両輪で解決

オフピーク需要

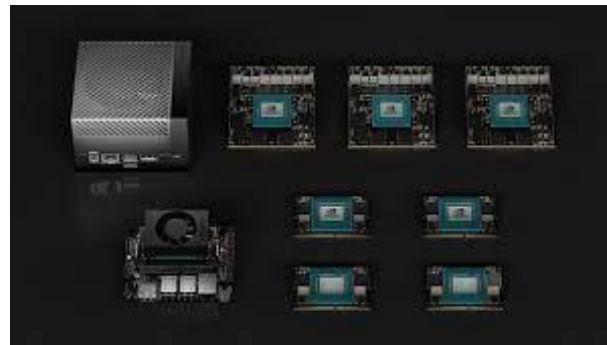
オフピーク需要の取り込み

- ・バッチ処理
 - 学習・解析
 - レポート・インテグレーション業務
- ・性能重視
 - 人間と同等のコスト
 - 人間以上の性能

分散学習

1分・1秒単位での学習の実現

- ・オーバーヘッドの改良
- ・分散学習



さいごに

さいごに

30社限定!
無料モニター
募集

東大発AIスタートアップ Lightblue

「法人向け 生成AI活用ワークショップ」 サービス提供開始

企業内での生成AIの導入・活用をサポート


LIGHTBLUE

