

人とAIとの共生社会の実現における 生成AIのインパクト

栗原 聡

慶應義塾大学 理工学部
慶應義塾大学 共生知能創発社会研究センター

自己紹介

慶大・大学院理工学研究科修了，博士〔工学〕
 慶大・共生知能創発社会研究センター・センター長
 大阪大学産業科学研究所・招聘教授
 人工知能学会・副会長／倫理委員会・委員長
 JSTさきがけ社会変革基盤・領域統括

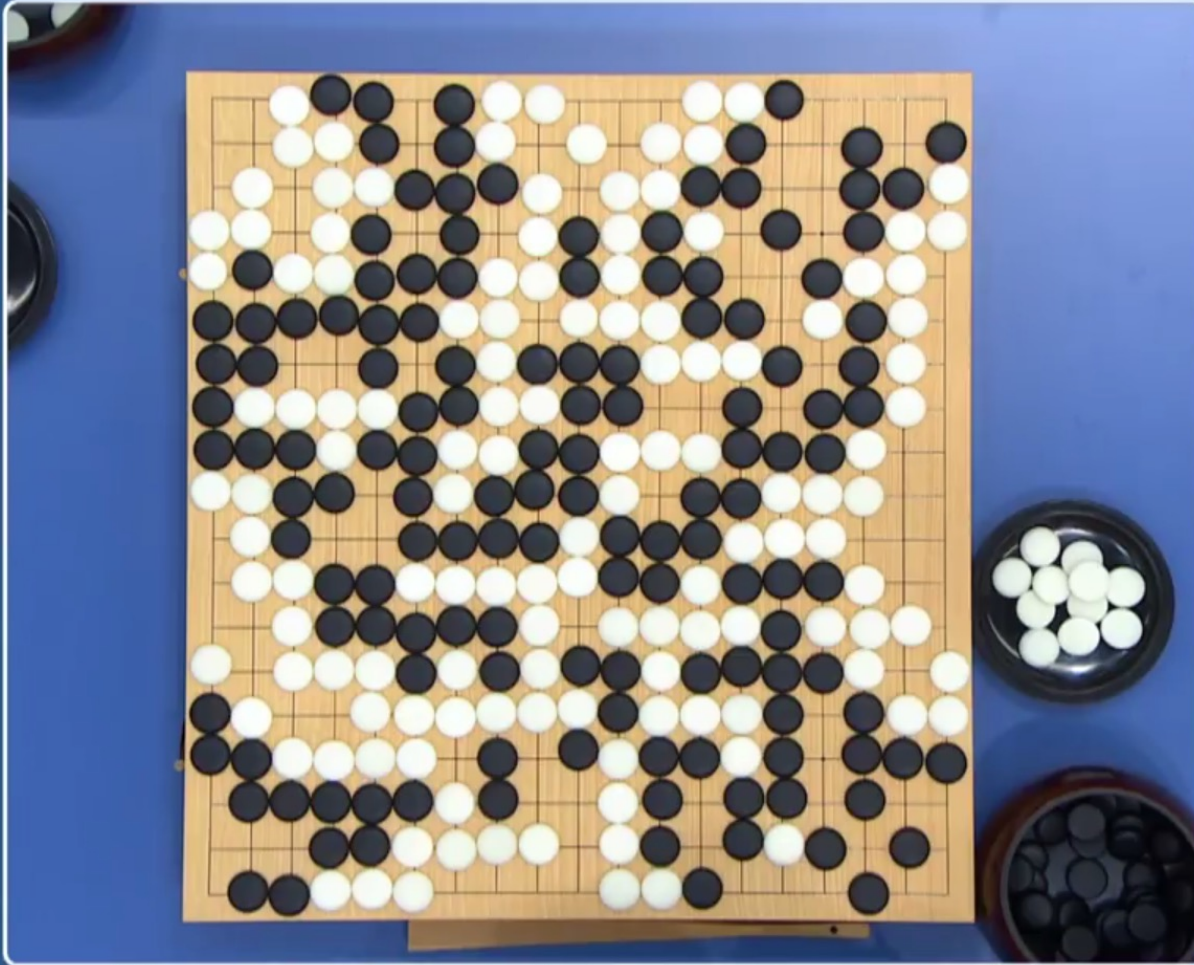


NTT研究所→大阪大学→電気通信大学を経て，2018年から現職．電通大に国立大初の人工知能先端研究センター（初代センター長）を設立．情報法制研究所上席研究員，総務省情報通信法学研究会構成員など．人工知能学会理事・編集長などを歴任．マルチエージェント，複雑ネットワーク科学，群知能，計算社会科学などの研究に従事．

※TEZUKA2020にて手塚治虫新作「ぱいどん」の制作でのAI担当．



碁盤の状況を画像として入力



 柯洁 KE JIE
00:13:28

 ALPHAGO
01:30:35



科学技術の加速度的進化

Deep Blue 1997
将棋終結宣言 2015
AlphaGo 2016
AlphaZERO 2017

強さ



時系列の学習

Deep Learningは
最初は 画像・音声認識で効果を発揮した。

もう一つ重要なデータが時系列データ。
言語, 感染推移, 株価, 自然環境, etc

膨大な言語データで単語の繋がり方を学習する

言語モデル→言葉を単語の出現確率でモデル化したもの

大量のテキストデータを使って事前に学習したベースモデル

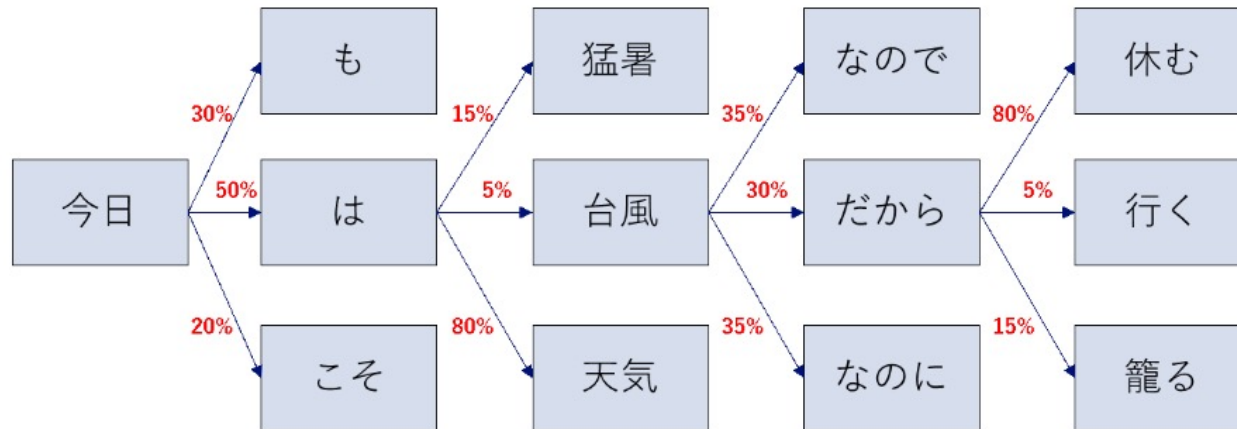
これまでは、ベースモデルに対する**ファインチューニング**で精度を高める必要があった。

GPT-3は膨大なテキストデータを用いて学習することでファインチューニングが不要→**AGI**.

Wikipedia やWebサイトから収集されたデータ等からの45TBものテキストデータを収集，前処理により570GBのデータセットを学習に用いている。

1750億個のパラメータを持つ自己回帰型言語モデル（ある単語の次に出てくる単語を予測するモデル）を学習した巨大な言語モデルを作成。

単語の出現確立でモデル化した言語モデルのイメージ



ChatGPTにとって文法は？

文法さえあればどのような言い回しにも対応出来る？

表, プログラム, 詩, シナリオ, ……
膨大なジャンルに対応するには??

VS.

確率のみでよい？

事実, 統計的な方法がかなりいけてしまった.
しかし. . . 突き抜けなかった.

スケールさせたらとんでもないことに！！

答えは 量・大きさ だった！

※スケールリング則

Model scale (training FLOPs)

<https://arxiv.org/pdf/2206.07682.pdf>

スケールさせたらとんでもないことに！！

—●— LaMDA —■— GPT-3 —◆— Gopher —▲— Chinchilla —◆— PaLM - - - Random

2022年07月25日 19時00分

ソフトウェア

「AIに意識が芽生えた」と主張したGoogleのエンジニアが解雇される



Googleの対話特化型AI「**LaMDA**」と対話して「LaMDAは人間と同じように意識を持った存在である」と主張し、Googleから休職勧告を受けていたGoogleのエンジニアが、ついに解雇されたことが明らかになりました。

Google Fires Blake Lemoine, Engineer Who Called Its AI Sentient

<https://bigtechnology.substack.com/p/google-fires-blake-lemoine-engineer>

Model scale (training FLOPs)

<https://arxiv.org/pdf/2206.07682.pdf>

スケールさせたら, とんでもないことに!!

—●— LaMDA —■— GPT-3 —◆— Gopher —▲— Chinchilla —◆— PaLM - - - Random

2022年07月25日 19時00分

ソフトウェア

「AIに意識が芽生えた」と主張したGoogleのエンジニアが解雇される



新しい技術の導入での変化はそれが目的

同じ技術でスケールしたことでの変化には気づかなかった



意識を持ったと実感してしまった

Google Fires Blake Lemoine, Engineer who Called its AI Sentient

<https://bigtechnology.substack.com/p/google-fires-blake-lemoine-engineer>

Model scale (training FLOPs)

<https://arxiv.org/pdf/2206.07682.pdf>

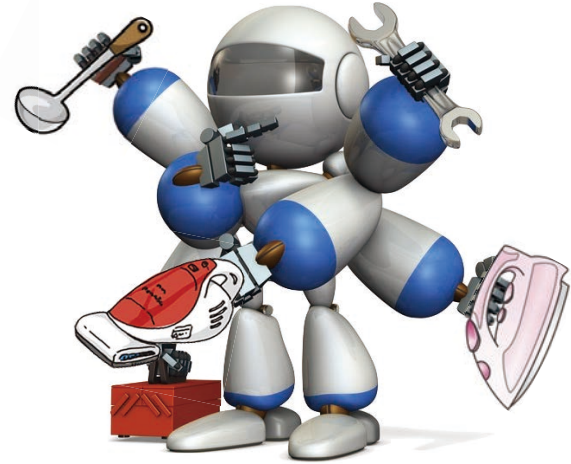
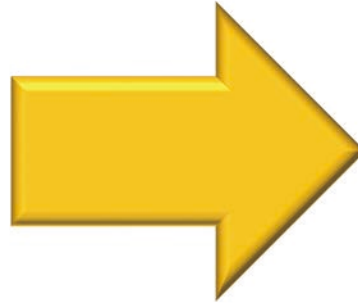
スケールさせることの重要性

小規模な試行では大規模導入した際の結果を予測することは難しい

スケールすることで性質の相転移的变化が起こる！

→複雑系ダイナミクスでは常識

AGI (汎用AI) とは？



用途限定型
弱いAI？



汎用型

道具としてのAIでの汎用AI→生成AI（まだ発展途上）→**AGI**
どのように個々の能力を利用するのは**人**が選択する。

自律型AIが汎用性を持つということ（まだ存在せず）→**自律型AI**
AIが状況を判断して能力を選択できる必要がある（**強いAI**）

※アトム, ドラえもん, HAL9000, T-800. . . .

強いAIは意識を持つ (John Rogers Searle)

人の脳でのダイナミクスとしての意識・感情が生成される同じ仕組みである必要はない。

- 社会性を必要とする人には必要な機能
- AIに必要とは限らない。
- 人と共生する場合に必要

あくまで人がAIにそう感じるのであればそれでよい (フェイク?) .



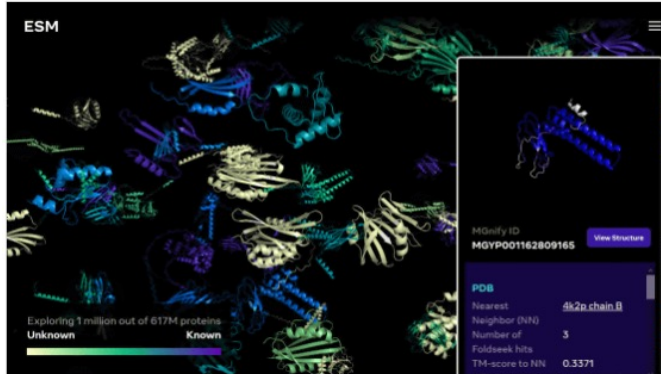
そう感じる仕組みを持つAIにおいて**AIならではの意識・感情**の仕組みをAIが持つと考えている

シンボルが並ぶ問題空間への適応

2022年11月02日 12時25分

サイエンス

6億種超のタンパク質構造を予測したAI「ESM-2」によるデータベースをMetaが公開、予測速度はAlphaFoldより60倍高速



FacebookやInstagramを運営するMetaが、タンパク質の立体構造を予測する言語モデル「ESM-2」を使用して6億1700万種以上の**メタゲノム**タンパク質の構造を予測したデータベース「**ESM Metagenomic Atlas**」を公開しました。

<https://gigazine.net/news/20221102-meta-esm-metagenomic-atlas-protein-structure/>

GenSLMs :ゲノムスケール言語モデルによるSARS-CoV-2の進化ダイナミクスの解明 GenSLMs: Genome-scale language models reveal SARS-CoV-2 evolutionary dynamics

Maxim Zvyagin, Alexander Brace, Kyle Hippe, Yuntian Deng, Bin Zhang, Cindy Orozco Bohorquez, Austin Clyde, Bharat Kale, Danilo Perez-Rivera, Heng Ma, Carla M. Mann, Michael Irvin, J. Gregory Pauloski, Logan Ward, Valerie Hayot, Murali Emani, Sam Foreman, Zhen Xie, Diangen Lin, Maulik Shukla, Weili Nie, Josh Romero, Christian Dallago, Arash Vahdat, Chaowei Xiao, Thomas Gibbs, Ian Foster, James J. Davis, Michael E. Papka, Thomas Brettin, Rick Stevens, Anima Anandkumar, Venkatram Vishwanath, Arvind Ramanathan
 bioRxiv Posted:October 11, 2022.

DOI:<https://doi.org/10.1101/2022.10.10.511571>

<https://medibio.tiisys.com/103670/>

まさにドラマ的な展開が.....

Google, Facebook も大規模言語モデルを開発していた。
公開するもネガティブな回答をしたり間違ふことが指摘され
炎上, 過去に公開を停止した経緯がある。

→公開に躊躇するようになった

OpenAIはスタートアップで非営利団体として発足
(イーロンマスクからの多額の投資にて研究を行っていた)

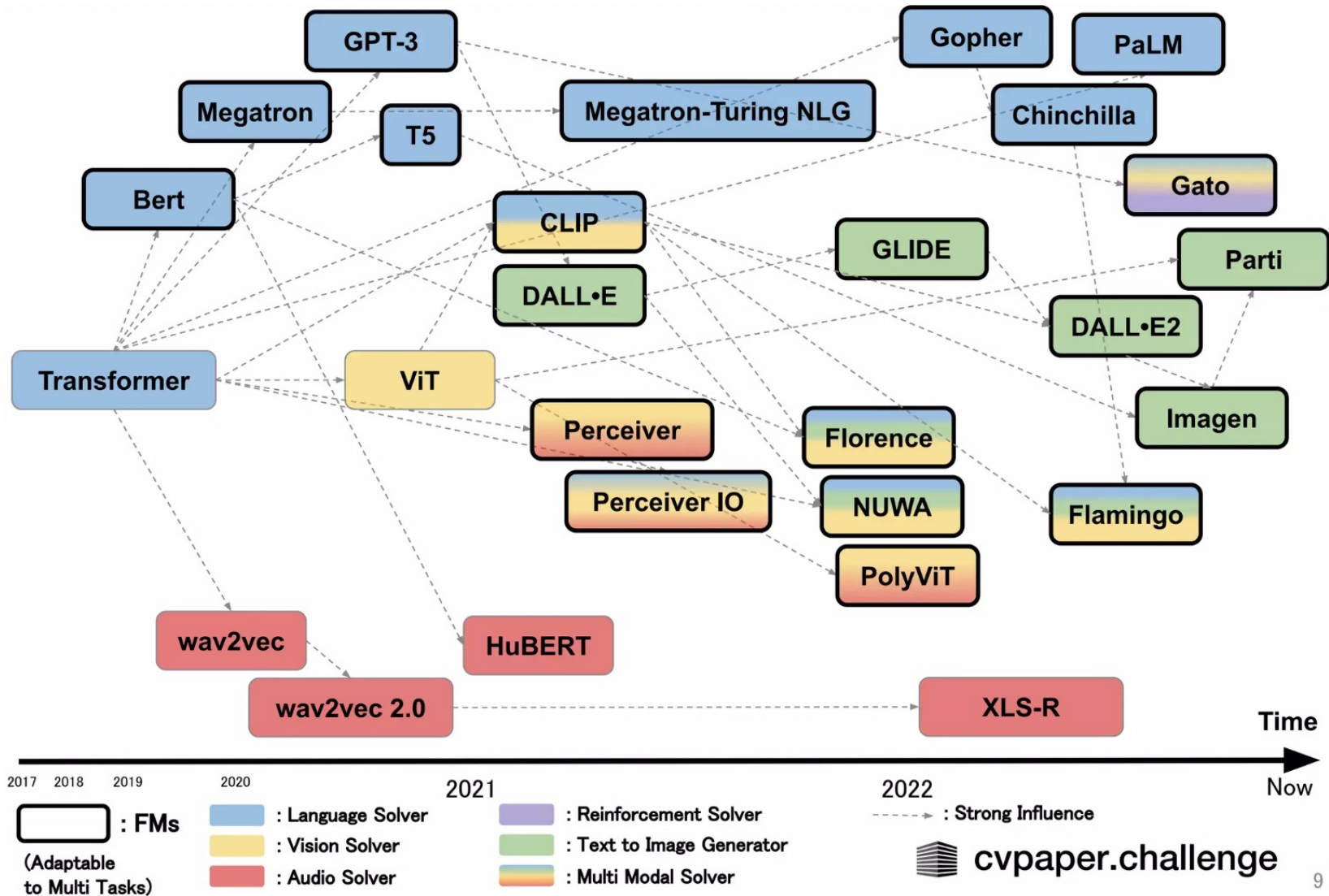
炎上からの学習→しっかりした対策を施す〔**AI Alignment**〕

満を持して公開, しかも無料で。

※イーロンマスクはその前に撤退

→この機を逃さぬように**MicroSoft**が投資！！

多くの生成AIが登場している



多くの生成AIが登場している



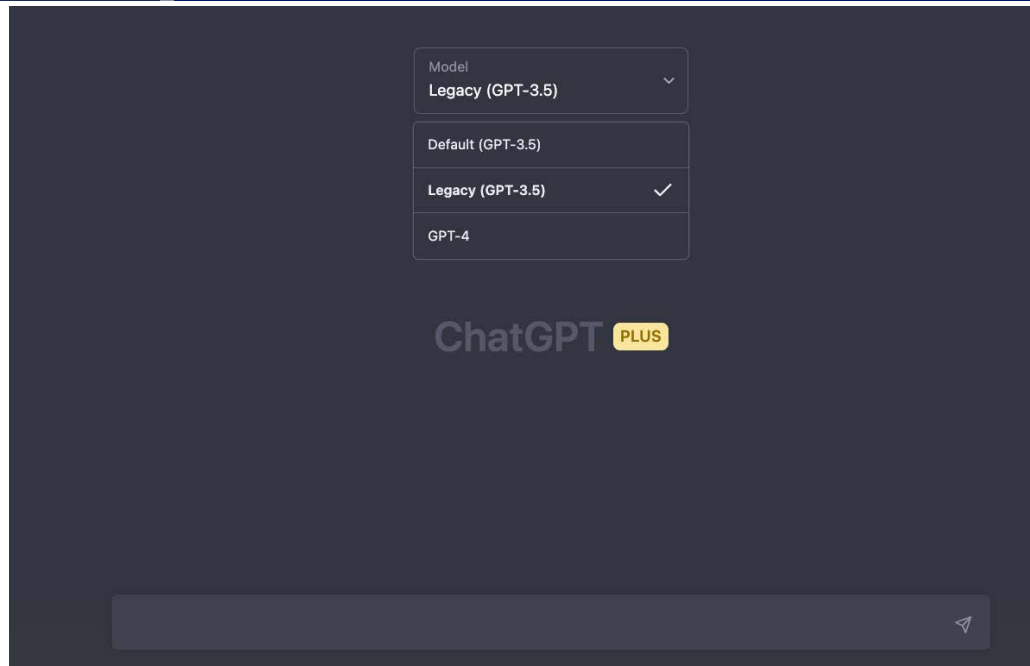
これまでの技術は徐々に社会に浸透

昨年からの生成系AIは**一夜にして世界を変える**
 これまでにない現象

ChatGPTで盛り上がり始めて矢先GPT-4が発表

加速的変化は危険！

ChatGPTとは？



ChatGPTは2022年11月に公開
すでに3段階のバージョンアップがされている。
最新がGPT-4

※GPT-3.5は 1750億パラメータ

GPT-4は1兆パラメータで**あろう**と言われている。

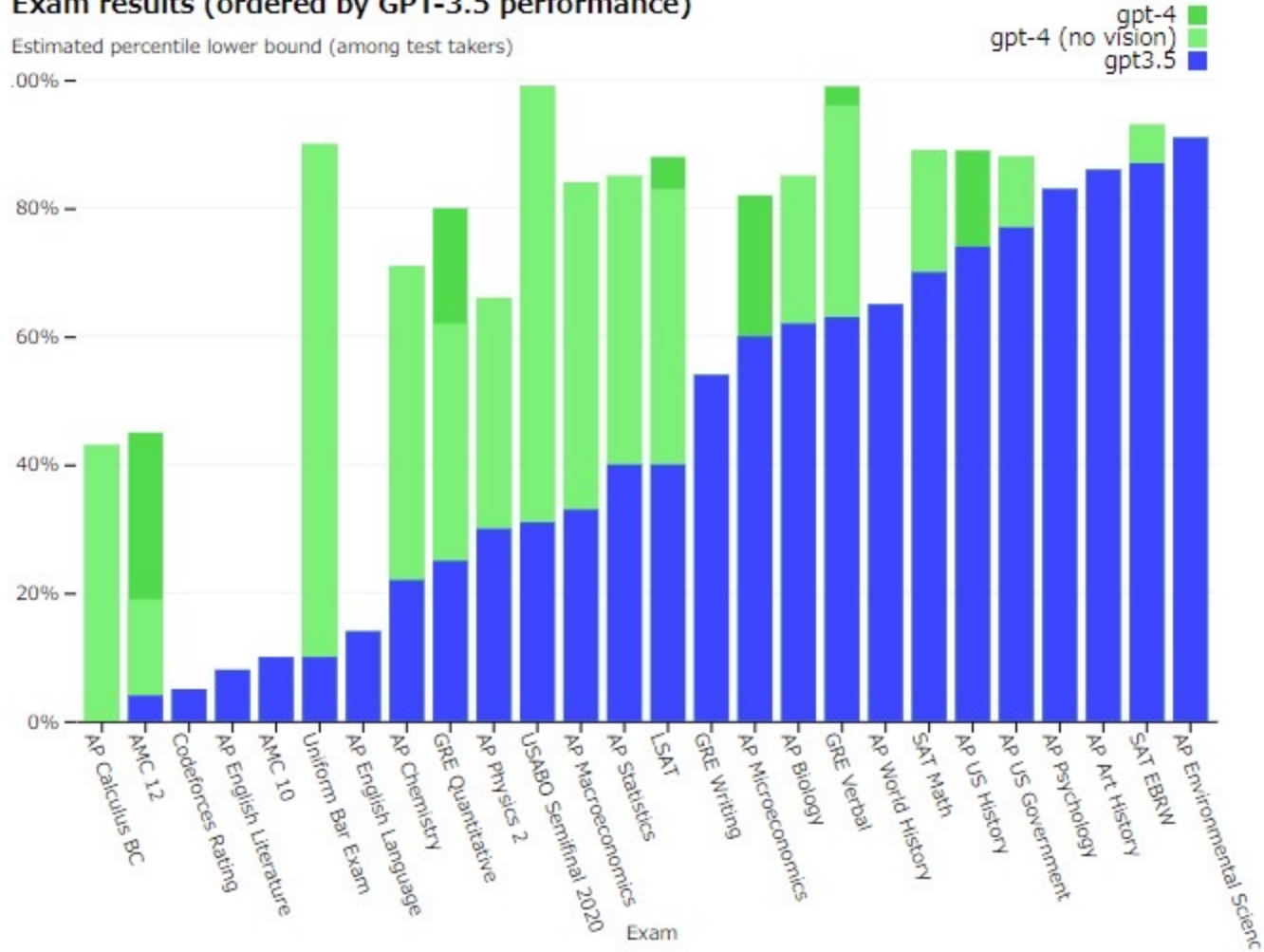


※OpenAIはそもそもオープンにすることが重要であると主張していた。
→GPT-4からは180度転換, オープンにすることは間違いであった

GPT-4 : 成績もすごい (学習データはChatGPTと同じ)

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)



ChatGPT〔対話AI〕は検索システムではない

Google 富士山と高尾山の比較

約 1,720,000 件 (0.62 秒)

累積標高差の比較では2.9倍の差が。高尾山を1単位として比較すれば、富士山の標高差は「2.9高尾山」です。 2020/03/17

登山ガイド.net
<https://登山ガイド.net> > 山情報 > 富士山

富士山は高尾山に比較して6倍難しい? ~いいえ - 登山ガイド.net

登山ガイド.net
<https://登山ガイド.net> > 2019 > May > 21

初めての富士登山の前に登るべき山① | 登山ガイド.net

2019/05/23 — 東京で最もメジャーな山、高尾山。標高は599m。登山口となる高尾山口駅の標高は約200m。高尾登山の標高差は単純計算で400mです。つまり富士山の標高...

富士山と高尾山の比較 の画像検索結果

ダイヤモンド富士 富士山登山 尾山山頂 日没 高尾山頂

標高差 (アール実績)	距離 (km) (ツアー実績)	実績 C T	気温 (hPa)
1,765	16.6	17:15	65
612	8.4	4:34	94
1,153	8.2	12:41	-29
2.9倍	2.0倍	3.8倍	-31

フィードバック

ChatGPT〔対話AI〕は検索システムではない



検索は膨大なWebサイトから関連するサイトをリストアップするだけ。

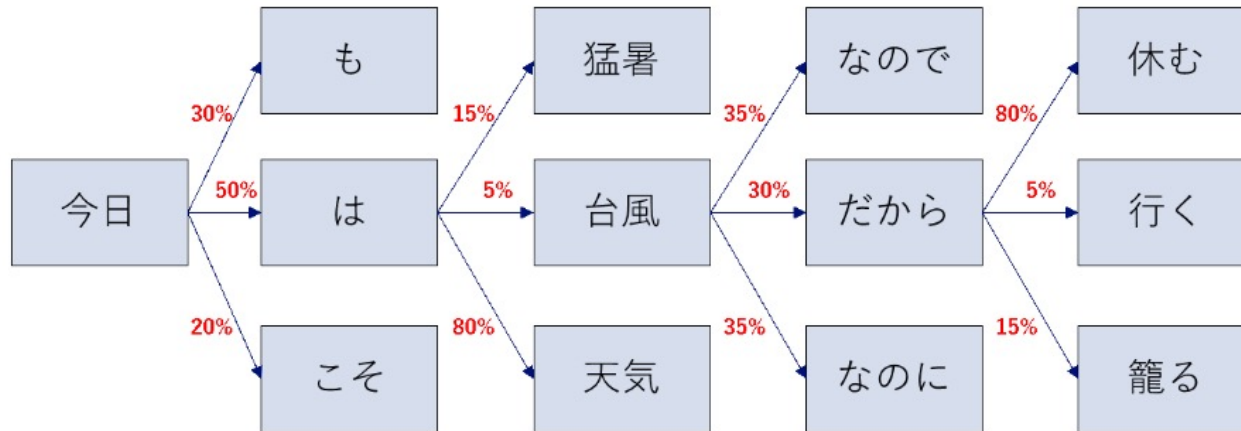


ユーザがアクセスして所望の情報をまとめる



ChatGPT〔対話AI〕がしていること

単語の出現確立でモデル化した言語モデルのイメージ



ユーザからの入力文に対して，モデル〔蓄えられた知識〕から最もそれっぽい回答を出力している〔それが凄い〕。

しかも対話形式に特化されている。

道具は最低限を使い方が分かってから使うもの (^^;;

ポイント1



エライ人「自分について間違った出力になるのですよ (笑)」
→ 「ChatGPTってダメだねえ」

✓ '栗原 聡' を検索しています

✓ 回答を生成しています...

栗原聡とは、慶應義塾大学理工学部の教授で、人工知能やプログラミング教育の研究をしています。[1](#) [2](#) [3](#)

栗原聡さんは、人と共に進化する次世代人工知能に関する技術開発事業や、NHKの解説委員室でプログラミング教育と考える力について語るなど、社会的な活動も行っています。[2](#) [3](#)

詳細情報: [1. st.keio.ac.jp](#) [2. k-ris.keio.ac.jp](#) [3. nhk.or.jp](#)

2 / 15 ●

✓ '栗原 聡' を検索しています

対話AIと検索を連携させればよい！！
先に検索し,その結果も対話AIへの入力として
しているらしい



検索の終わり?? ググるの終焉??



シェイン・グウ
@shanejg

ChatGPTチーム技術幹部

日本語は英語より悪いですが、皆さんの役に立ったChatGPT使用例があれば #チャトル でお知らせお願いします。

決して、「#チャトル を流行らせたい」とか、「#ググる から #チャトル へ」とか、「#チャトル で流行語大賞狙ってる」とかやましい気持ちはないので安心して下さい。

Googleがコードレッドを宣言！

2022年12月23日 12時30分

ネットサービス

ChatGPTのリリースでGoogleは「コードレッド」を宣言、AIチャットボットが検索ビジネスにもたらす脅威に対応するためにチームを再割り当て



AI開発団体のOpenAIが発表した「**ChatGPT**」は、自然言語処理モデルの「**GPT-3**」の進化系である「GPT-3.5」がベースになっており、質問文を打ち込むと人間が書いた文章と見分けが付かないほどに高精度の文章で回答してくれます。そんなChatGPTの登場に、ユーザー数世界最大を誇る検索エンジンを持つGoogleの経営陣が事業に対する深刻な脅威への警戒を示して「コード・レッド」を宣言したと報じられています。

<https://gigazine.net/news/20221223-google-code-red-against-chatgpt/>

精度は向上するが完璧にはならない

Model: Legacy (GPT-3.5)

SA 慶應の栗原聡とは？



栗原 聡 (Kurihara Satoshi) は、日本の数学者です。慶應義塾大学教授として知られています。主に代
数学や数論に関する研究を行っています。

Model: GPT-4

SA 慶應大学の栗原聡とは？



栗原聡 (Satoshi Kurihara) は、慶應義塾大学 (Keio University) の教授であり、人工知能
(AI)、ロボティクス、計算社会科学、ゲーム理論などの分野で研究を行っています。彼は、これらの分野
でのアプリケーションや研究開発において、多くの貢献をしており、国際的な学術会議での発表や論文の執
筆を通じて、その知識を広く共有しています。栗原聡教授の研究は、情報科学や計算機科学に関連する
多くの分野にまたがっており、彼の貢献によって、慶應義塾大学の研究力が向上しています。

- 流暢な言い回しで返してくるが、鵜呑みにしないこと
- 検索的利用よりも**文章生成のための利用**の方が好ましい

ChatGPTの登場が意味すること

これまでのAIは特別な技術を持つ者しか利用できなかった。
ChatGPTは文章として入力することで最新AIを誰もが利用できる。



AIの民主化

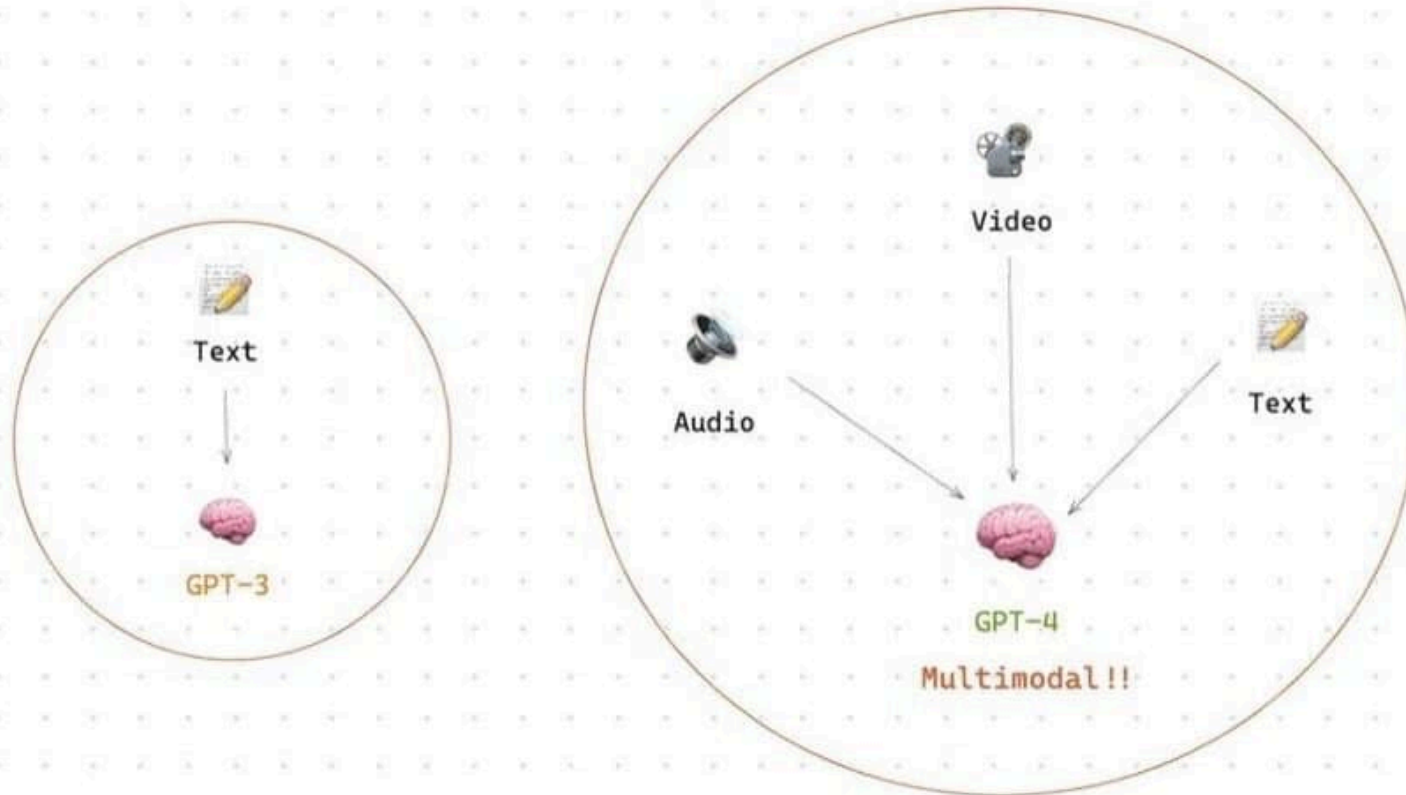
特別な技術を持たない一般ユーザでも最新AIを活用できることの意味

産業革命→インターネット→生成AI

という大きな変化

マルチモーダル化

GPT-3 vs GPT-4



https://scontent-mad1-1.xx.fbcdn.net/v/t39.30808-6/336135770_1901973913517289_2619615418300016163_n.jpg?_nc_cat=105&ccb=1-7&_nc_sid=5cd70e&_nc_ohc=Th2tri8SAvAAX9m-Cw_&_nc_ht=scontent-mad1-1.xx&oh=00_AfDj1x48PX2I5ppOYDtXrW4DaSveeBC6zrPTYjRNDrgAVQ&oe=641789A2

AIに創造は可能か？

そもそも創造するとはどういうことか？

創造性とは？

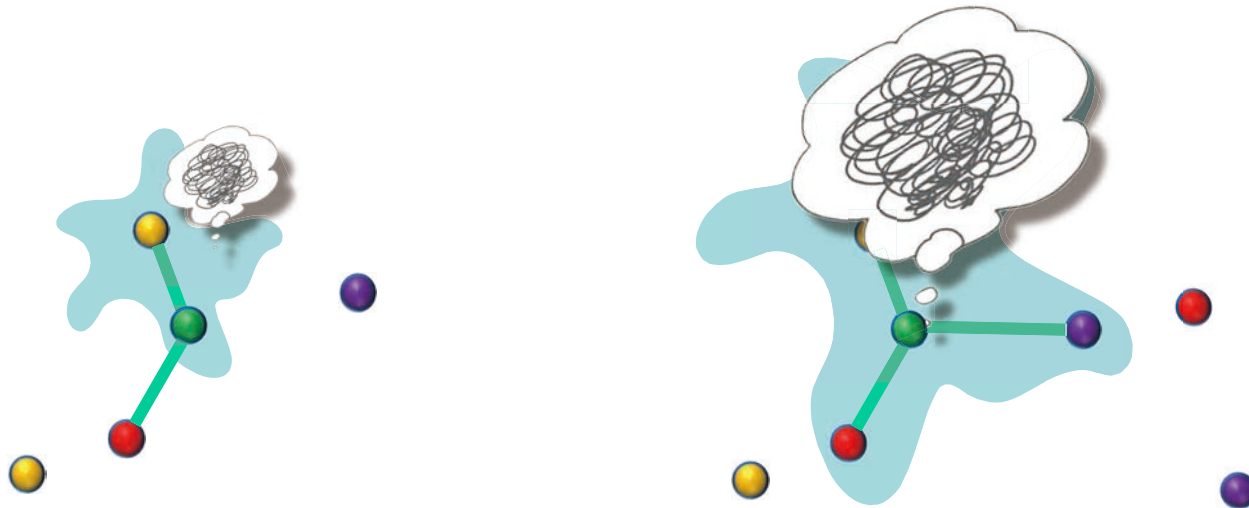
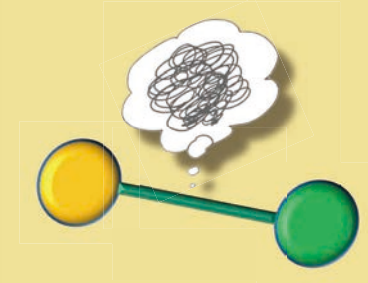
点の発見
(クリエイション)



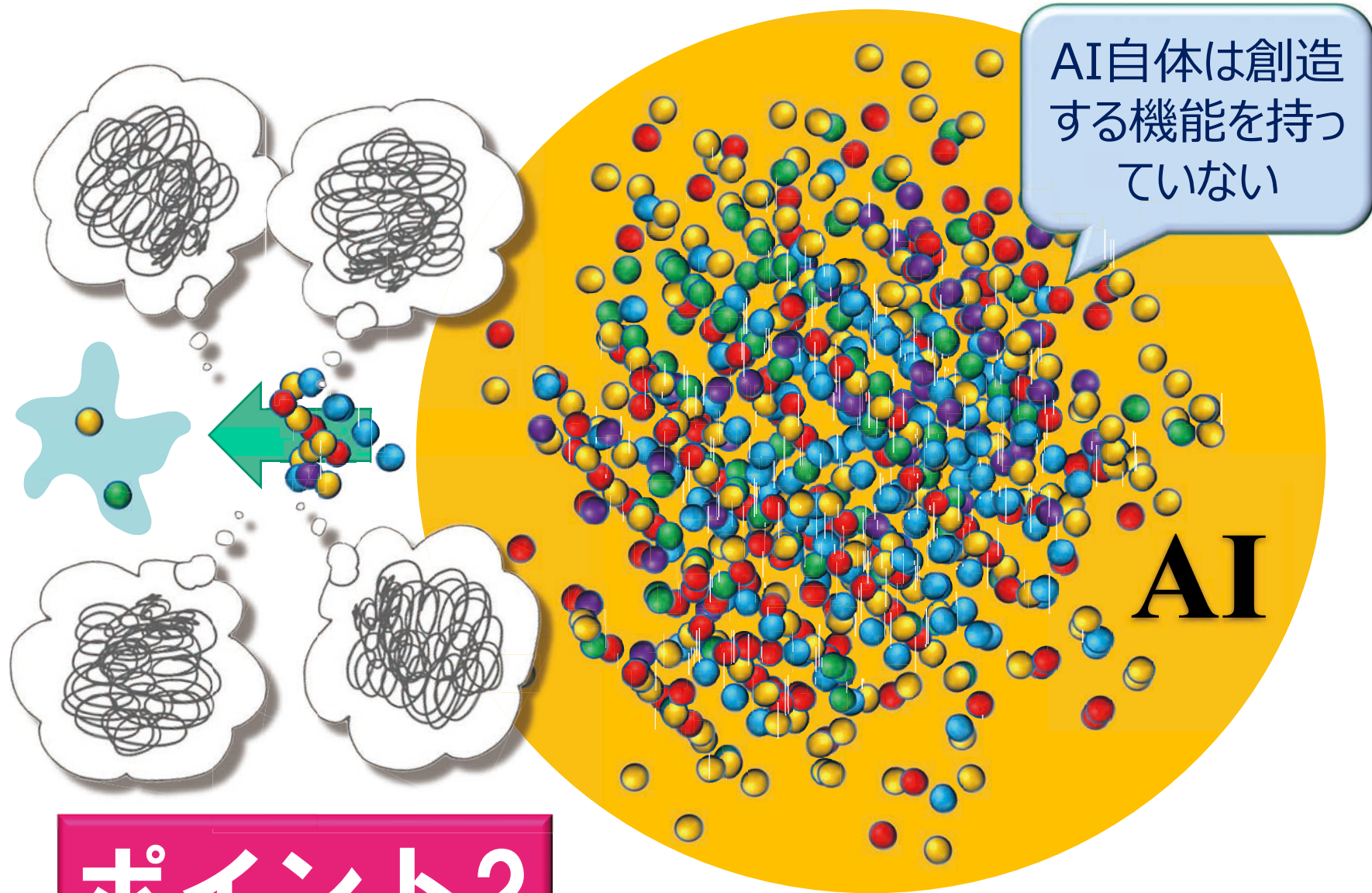
※偶発性

点と点を結ぶ繋がり
(イノベーション)

ネットワークによる創発
＜妄想力＞



創造を効率的にサポートするAIの立ち位置

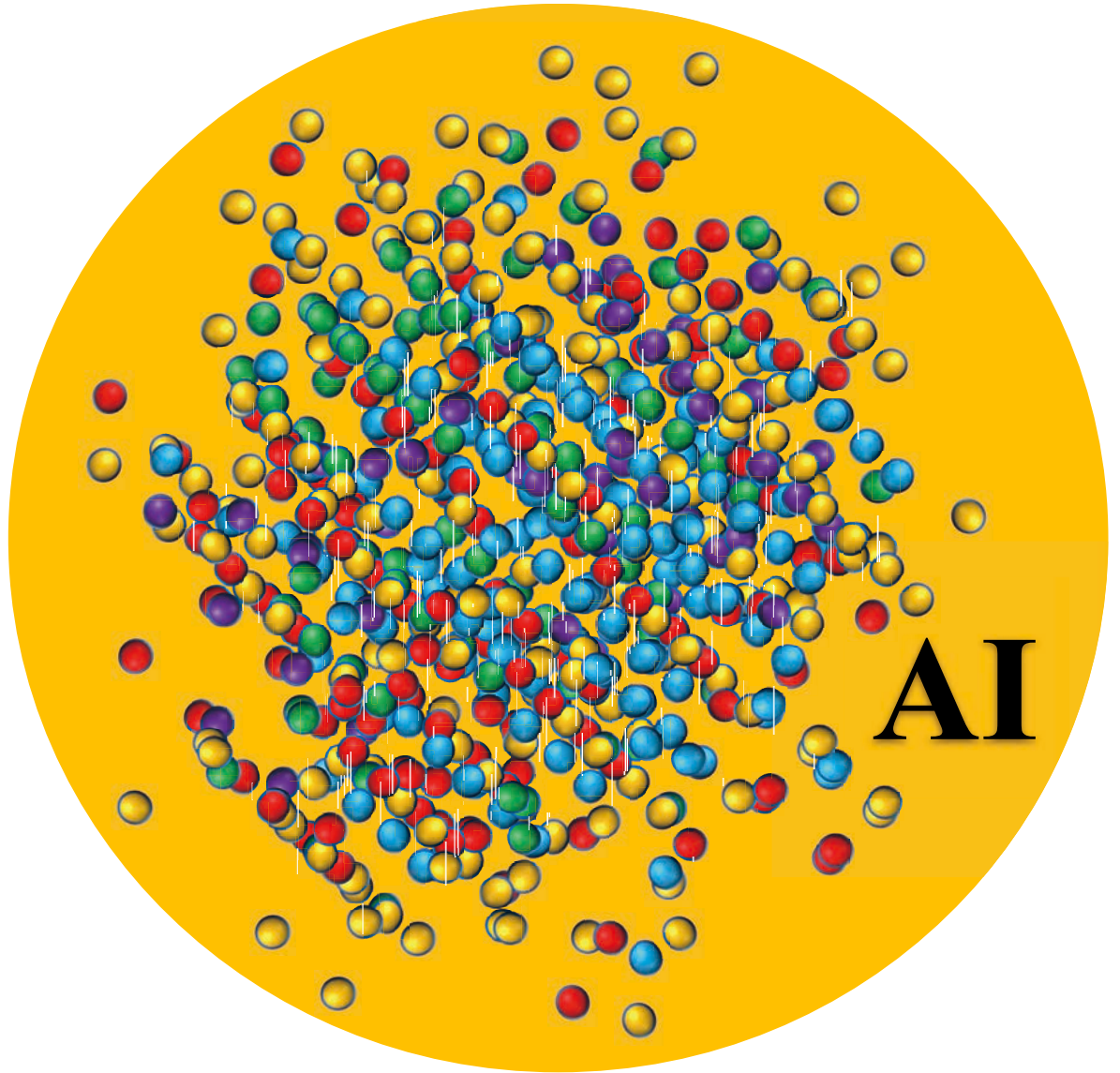
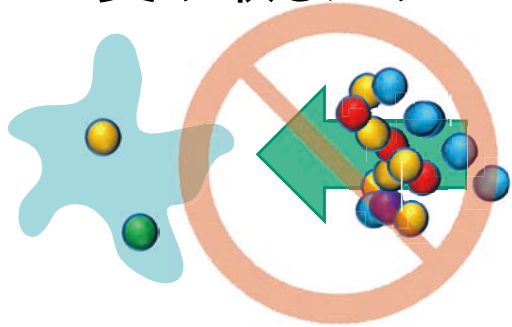


ポイント2

ただし.....ただ受け取りだけなら思考力の低下を招く



受け取るだけ...



生成AIは日本経済浮揚のための起爆材となるか？

YES 具体的に浮揚させる方法は？

創造サポートツールとしての使用
圧倒的な案の提示・ブレスト相手
※人同士よりもはるかに多くの示唆を得られる。

効率化ツールとしての使用
メール, 報告書, 仕様書, 要約, 情報整理
※大量の文書を読み込んでの作業 (AIならではの)

**誰もがGPT-4 (それ以上の) レベルの
大規模生成モデルを利活用できる環境の構築 (国策として)**
※マルチモーダル / System2 / 高自律・高適応能力, etc

※富岳, Spring8

お互いに長所・短所が異なる

AI

人

計算
記憶

対話

推論, 仮説検証,
論理, 状況理解,
他者理解, 因果,
時間の把握……

我々はどうすればよい？



- なぜだろう？（好奇心を持つ）
- 五感を育む
- 行間を読む能力
- 自己の意見を論理的に主張する
- 答えよりプロセス（問題から計算式を導出するまでの過程）
- 複数の方法を考える
- 別の角度から見る
- 俯瞰できる（メタ）能力

ポイント3

適応力・文脈理解・多様性・感性・社会性

※一回りしてあらためて人間力が必要に！

進む格差

残念ながらAIを使える層と使われる層の格差が進む

使える層は極めて少数→利益獲得

圧倒的多数の使われる層（思考力の低下）

すぐにでも出来ること

社会インフラへの導入を急ぐべき〔行政コスト削減〕

等しくAIの恩恵を分配／ベーシックインカム的な取り組み

教育現場での効率化→生徒に向かう時間の確保

ホモ・デウス ～人も変わる～

ホモ・デウス〔神〕は ホモ・サピエンスの後継種。
生物的進化ではなくテクノロジーによる進化として出現する。



健全者(富裕層)によるアップグレード



可能性:誰もが使えるということが意味すること

これまではAIから距離のあった人が自身の創造力を発揮できるということ。

想像できないいろいろな可能性が埋もれている。

※例えば高齢者層が利用することでの社会の変化は絶大

いかにしてAIにリーチさせるか次第



2023年04月21日 15時52分

ソフトウェア

ユーザーの同意なしに訓練した可能性のあるChatGPTの「OpenAI」が厳格なプライバシー法を重視するEUで法的な問題に直面、「規則の準拠は不可能に近い」と専門家



プロンプトから生成されたコンテンツは著作権保護対象外。

コンテンツへの人の手入れによる新たなコンテンツ生成は対象？

人間レベルの文章を生成するChatGPTがEUのデータ保護規則に違反している疑いで、開発元のOpenAIがイタリアのデータ保護機関から調査を受けています。当局は2023年4月30日までにイタリア及びEUの法律を順守するよう呼びかけていますが、AIの専門家は「このタスクは不可能に近い」とし、ChatGPTおよびOpenAIが抱える問題について詳述しました。

OpenAI has until April 30 to comply with EU laws — ‘Next to impossible,’ say experts

<https://cointelegraph.com/news/openai-has-until-april-30-to-comply-with-eu-laws-next-to-impossible-say-experts>

EUの一般データ保護規則(GDPR)の下では、技術系企業は個人データを使って製品を訓練する際にユーザーの同意を得なければなりません。さらに、EUで事業を展開する企業は、データの収集と共有をEUの人々が拒否できるよう選択肢を与えなければなりません。

https://gigazine.net/news/20230421-openai-eu-laws-problems/?fbclid=IwAR1BwKvCBKKPLY1iS11FPikNDMfBE-aG7LDTndY5p_qi3MVIUBYFLqxmBCw

Arts

AI生成画像が写真コンテスト入賞、出品のアーティストが受賞辞退

2023.04.21 Fri posted at 18:15 JST



コンテストで入賞を果たした、AI生成画像で制作した写真作品/Boris Eldagsen

日本の立ち位置

完全に蚊帳の外の状態

LLMによる研究・サービス創発の可能性はあるものの、
LLMという基盤において単なるユーザという立ち位置。

→データ整備， 計算リソース， 人的リソースの全てで弱い

国産の必要性の声はあるも， 今のタイミングを逃すともはや
追いつくことは不可能→開発しても無駄か？？

ユーザとして割り切ったのLLM活用という路線もありえるが，
経済安全保障的にはまずい。

→やはり国産の必要性はある。

→小粒でも大規模同程度の性能が発揮できる技術の熟成



※2つの展開
(良い方向と悪い方向)

G7群馬高崎デジタル・技術大臣会合での「信頼できるAI」(1/2)

．．．．民主主義的な価値観に基づく「信頼できるAI」の普及を目指すという大枠では各国が一致し、
▽国際的な技術の標準化を進めることになったほか、
▽国際機関で課題を分析することなどを盛り込んだ行動計画が採択されました。

ただ、総論では一致しても、各国の姿勢の違いを超えて、どう具体的なルール作りを進めていくかは今後の課題です。．．．．

(NHK)

<https://www3.nhk.or.jp/news/html/20230430/k10014054401000.html>

G7群馬高崎デジタル・技術大臣会合での「信頼できるAI」(2/2)

- EU,日本というAI劣勢チーム 対 米中AI優位チーム という構図
- 前進しようという姿勢を出すことが出来たのは安心した。
→規制を厳しくすると開発が止まり、地下に潜ってしまう
- **具体的なことはすべてこれから**→新しい斬新な提言まではさすがに無理だったのでは。
- 社会(作る側使う側)は勝手に突き進むが、その動きに臨機応変に合わせることも重要だが、しっかりした統一した指針を示すことは必要。
- AI利用においては、ありきたりではあるが、使いながら軌道修正しかない。
失敗に対する具体的な対応をどうするの？
- 信頼できるAIとするための基礎研究推進のためのしっかりした研究開発体制の拡充をもっと主張すべき。
- AI先進国の中国が入っていない。実現性の懸念あり
- そもそも信頼とは？ 単に制御可能性・透明性のみを指すのではないはず。
学習に利用するデータへの保護は難しい。日本のやり方は現実的。
生成されたモノが似ることへの保護はまだ議論が必要。
- 地域によって使えるAIが異なるような落とし所→自律型AIは日本のみで発展??
- 一極集中という現実 (GAFAM) を踏まえた対応が必要。
→データ流通させればなんとかなるってもんじゃない。

課金してのAPI利用(Google, MS等のサービス利用と同じ)
※FMは国内代替サービスがない。 **経済安全保障**。

ドメイン限定モデルの開発・運用
※大規模であることに超したことはない。
※小粒を足しても大粒に及ばない

追いつき



次世代AIの実現
※大規模基盤モデルの上での実装である必要。
※日本型AIの実現

リードする

ホモ・サピエンスの自助では対応不可能な状況

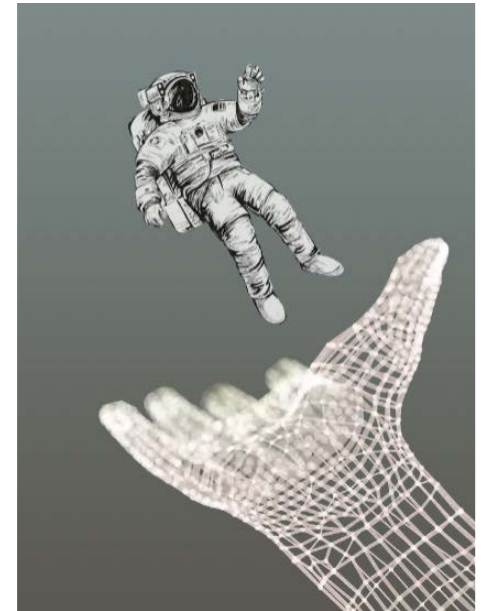
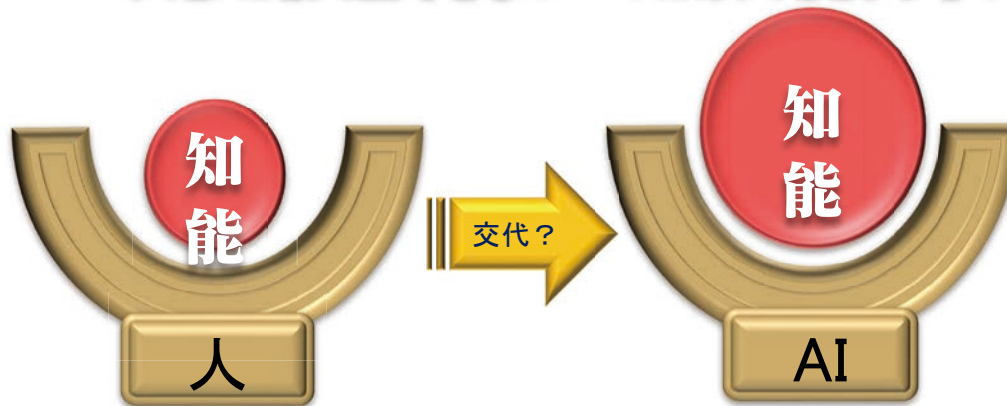
AIに思考を操作される。間違う情報に混乱させられる……

そもそも、人がSystem1化しつつあることこそが問題！！

根源→SNSがもたらす混乱（**デマ・フェイク・フィルターバブル……**）
※AIはその混乱を増幅させてしまう役割

救うのもAI（自律型AI）

人とAIと共生できる？ AIは人をどうする？



ポイント5

現在を生きる我々次第

