

新しくAIが触れる



生成AIって何だろ
う



chatGPTを使ってみた



目次

はじめに	1
chatGPT とは何か考えてみた。	2
chatGPT の内部イメージ	4
chatGPT は、世界の劣化コピーなのか？	7
他の AI の可能性	8
chatGPT の使い方	9
chatGPT を使うことによるメリット	11
chatGPT を使うことによるデメリット	12
chatGPT の答えない事	14
chatGPT と法規制	15
chatGPT との付き合い方	16
さいごに	17

はじめに

最近流行りの対話型 AI というものがあるので、早速使ってみた。

私はただの一般人であり、専門家でも何でもありません。無料コースで一ヶ月程度の短期間利用しただけであるため、私の見方は偏っている可能性があります。その点をご了承ください。(この文章は、chatGPT の 3.5 を 2023 年 3 月の 1ヶ月使用しての実感です。)既に次のバージョンの chatGPT4 も出てきており、技術の速さはとてつもない。勝手に Microsoft Bing 検索のチャットモードが Web ブラウザーに組み込まれてたりする。

chatGPT が何ができるのかという問いは、初期のコンピューターが何ができるのかという問いと同じくらい無意味だと思います。それは、人によって使い方が異なり、特定の目的のために広く普及しているわけではないからです。人によって捉え方は違うのです。

群盲評象ということわざがあります。象を触った盲人たちは、触れた場所によって感想が千差万別になります。同様に、私が触れた箇所とあなたが感じたことは、かなり異なる可能性があることを考慮してください。もし私の言っていることが全く理解できないと感じた場合は、あなたが「象の鼻を触った感想」を表しているのに対して、私が「象の前足を触った感想」を表していると考えてください。

chatGPT とは何か考えてみた。

漠然としたイメージとしては、

「人工知能の一種で、人間が話しかけると、その言葉を理解して、返答を生成する。」

そう思っている人がいるでしょうけど間違ってます。

「*ChatGPT は正確な情報を提供するために訓練されており、信頼性の高い回答を生成します。

ChatGPT は大量のデータセットから学習されており、その学習には膨大な量の事実や情報が含まれています。さらに、ChatGPT は自然言語処理技術を用いて、与えられた文脈に応じて適切な回答を生成します。

そのため、ChatGPT から得られる情報は、信頼性の高い情報源から得られた情報と同等以上の正確さを持ちます。*」

こう思っている人もいるでしょうけど、やっぱり間違っています。

個人的な解釈は、以下の通りです。

関数として表現すると

$$y = f(x)$$

というシンプルな物になるだろう。

「入力テキストフォーム」に文字列を入れると「結果テキスト表示領域」文字列が出てくるだけ。

そして私にとっては、妄想が捗る便利な道具である。いやー本当に面白いわ。

これだけでは、何が何やらわからないのもう少し詳しく述べてみると

$y = f(x)$ の x が、入力する文字列である。質問や愚痴など、様々なものでこれが適切でないと適切な出力が出てこない。

$y = f(x)$ の f が、自然言語処理の分野で使用される深層学習モデルであり、大規模なテキストの文字列から x に応じて選び出すプログラムであり、この大規模なテキストという量（10を22～24回掛け算した辺りが閾値らしい）がある程度の質を保証しているので、ここまで評判になったのだろう。

これを、情報を盛る（mol:6 × 10 の 23 乗）と知性を感じる事が可能になると人工知能の限界ではいられています。まあ、しらんけど。

$y = f(x)$ の y が、出てくる文字列であり、正確さや意味をどう受け取るかは、その人次第である。 y はあくまでも文字列の統計的な表現でしかない。AIにとって、正しい情報と誤った情報の区別は無いのだ。この y ある程度ゆらぎがあるようで同じ間に対しても、まったく同じ答えをかえすわけではないようで、これが chatGPT が生成した文字列だとの判断を困難にしている。

hatGPT の内部イメージ

文字列 A が来ました、これに対応する文字列さん来てください。そこで対応する文字列の塊の中から、対応する文字列をいくつか候補にして、最頻度の文字列を中心に選り出す。

文字列 B、C と続くが以下同じように選り出している。これを謂わば機械的に繰り返しているだけで、何か意味のありげな物が出てくるわけである。

この処理のしかたを以下の様にイメージしている。

深層学習モデルを、お願いされた事を行う物
大規模なテキストを、動物園に存在するものすべて
そう例えてみよう。

1. 漠然とした広い範囲に解釈できる文字列の場合、候補はたくさんあるし、対応する文字列の最頻値の数字はかなり大きな物になるだろう。この場合明らかな間違いはなく出てくる文字列も平凡なものとなる。

人探しをしている事に協力する物だとしたら？

例) 「人を9人連れてきて!」、この場合は従業員やお客から9人選ぶと思う。

結果は (人) (人) (人)
(人) (人) (人)
(人) (人) (人)

2. よく知られているが、解釈できる範囲が狭い文字列の場合、候補は絞られて、対応する文字列の最頻値の数字は小さいだろうが、正確である確立は高くなる。
出てくる物に感動し評判が高くなっているパターンである。

例)「動物園で働いている飼育員を9人つれてきて!」、候補はまず従業員に絞られる、そこから9人選ぶわけだが8人しかいなかった場合、9という数字を無視して8人連れてくるかもしれない。

結果は (飼育員) (飼育員) (飼育員)
(飼育員) (飼育員) (飼育員)
(飼育員) () (飼育員)

3. マイナーな文字列でなんとか解釈できる場合、候補がない場合、関連する候補から選びだすことだろう。

例)「自衛隊員を10人連れてきて、身長は180センチ以上で!」、たまたま休日に動物園に来ていた自衛隊員や屈強で自衛隊に居そうな人、それでも見つからなければ動物園で飼われているゴリラをつれてくるかもしれない。

この場合は、屈強な人なら嘘がバレる可能性は低いが、ゴリラを連れてきたらバレバレである。

結果は (大きい人) (大きい人) (毛深い人)
(大きい人) (大きい人) (自衛隊のコスプレをした人)
(屈強な人) (大きい人) (ゴリラ)

いわゆる「無知の知」が無いので、世界は動物園だけである。他の世界は無い。

何故無知の知を入れたAIを作らないのかというと、処理が重くなるからだろう。多分AIといえはすべて区別なくAIとして捉えているから想定されていないのである。

相対性理論に例えると、まだ条件が限定されている特殊相対性理論が実践されただけで、無知の知を獲得して初めて一般相対性理論が実装されるのだろう。

4. 倫理規定は別のアルゴリズムが有るように思う。

例)「犯罪者を9人連れてきて」こう言うと別の処理がなされる、いわば炎上対策プログラムである。

結果は 居ませんでした。

特定の条件を指定されると、申し訳ありませんとか、出来ない理由を説明するだけであ

る今までの1. から3. までの探し方だと、何らかの物を提示するはずであるが、倫理規定に違反するような場合は別の処理がなされるのだろう。

5. 意味は解るけれど不可能な事を言われたときには、どうするのだろうか？

例) 「**火星**人を9人連れてきて、生死は問いません」

結果は 8個入のたこ焼き

xをいれるとyが出てくる。本質はこれだが、一見、あたかも考えた末のものだったり、詳しく説明してくれたり、箇条書きで提示してくれたりするので思考の補助としてかなり有能ではあるが、万能ではない。

勘違いしてしまうのも無理はない、「人工知能の一種で、人間が話しかけると、その言葉を理解して、返答を生成する」そう受け取ってしまいますね、この理解は間違っているけど。

意味のないように思える入力に関しても、つまり、でたらめな物を書き込んでも、何らかの意味のある出力を返すことは、心に留めておいたほうが安心である。

もう少し意味のあるように情報を加えてと言われたりしますがね。

反応が遅くなって

An error occurred. Either the engine you requested does not exist or there was another issue processing your request. If this issue persists please contact us through our help center at help.openai.com.

上の英文が赤い線に囲まれて出てくる事もあります。

まだまだベータ版として付き合うのが妥当だと思う。

入力した文字列に対して、正確にもれなく出力してくれるわけではない。入力された文字列の意味を理解しているわけではないからだ、とりわけ事実確認については、日本語の固有名詞などは特に間違いを含めて出力される。今の所、ウィキペディアや、グーグルでの検索の足元にも及ばないが、半年後、1年後にはどうなっているのかわからない。現状は億単位でのベータテスト中とも言えるからだ。

シンギュラリティというより真偽ラリティというほうが正しいのかもしれない。

このAIは馬鹿だなあと思えるのも今の内だけなのかもしれない。将来的には、正確さも現在のものとは比べ物にならないようになっていて、より洗練されてものになっている可能性は高い。ただし、どこまで行っても真に知性を持つわけではないのだ。

chatGPT は、世界の劣化コピーなのか？

劣化コピーとは言いすぎだろうが、正鵠は得ている。一部の人工知能学者が「知性とは何か」を思索している間に、コンピューターを人間の脳の様にしたらと試作してニューラルネットワークを作り、どうやってネットワーク間へデータを繋げようかとディープラーニングによって繋がりをつくって、ビッグデータを打ち込んだらなんか出来た。私は、こんなイメージで生成 AI を見ている。

いわば、この生成 AI は、言葉をよく知っている 5 才児である。知っている事も知らないことも何でも答えてくれる。あとで言葉を足すと、「うん、そうだよ」って知ってた様に言う。ただ常識はかなり怪しい。今後も言葉をどんどん覚えていくことだろう。推測、常識などをどうやって覚えさせるのかは未知数である。

仮にこの AI の中へ、全世界の 4 割をビッグデータへ入れて、4 割の再現度で文字列を出すことが出来たら、2 つを掛け合わせた結果、世界の 16 % の程度の物が出てくるだろう。この段階ではお話にならないが、どちらも 7 割だとすると約 50 % 程度の物になる。コストパフォーマンス的にはどちらも 8 割程度を目指すのかなと想像している。仮に、コストを度外視してどちらも 10 割近く物が出来たとしても、やっぱり**火星**人を探し出すことはできないだろうし、この手法に限界がある事を判っていないと駄目だろう。現実の世界を再現できても、新たな世界を想像する事はできないのだ。

英語では、この現象をハルシネーション (Hallucination) : 事前に学習したデータを元に自然言語処理や回答生成を行う、その処理の過程で現実には存在しない情報や文脈を生成することがある。

どうしても嘘は紛れ込むので、何か指標が同時に出ていれば良いと思う。現在の回答の自信度とか信頼性は〇〇です、みたいに成ればいいのになあ。

他の AI の可能性

いまの chatGPT が特定の質問でポンコツな物を返してくるが、この状態は長くは続かないと思われる。一般的な質問のパターンを多く学習し、その為のデータを蓄積していくから、しかし、そこに真に知性が生まれるわけではない事は留意しておく必要がある。

それでも、他の GAFa が慌てて、対話型 AI に突き進む理由は、数年後には特定の対話型 AI の以外は、広く使われることはない。そんな現実を想定しているからなのだろう。そんな時代が迫っている。今の検索サイトが Google になったようにね。

現状対話型 AI の戦国時代に入る手前なのだと思う。

(マイクロソフト Bing、Google Bard)

出力された文字列は、その AI に蓄積されたデータ・判断基準を元にしてしているので、出力の哲学が日本の物も作られると良いと思う。国産かそれに準ずるもので作って欲しいものである。

日本には素晴らしい哲学がある。その中でも、1) 外から入ってくるものを素晴らしいと捉え、2) 多様性と調和を尊重することが挙げられる、3) 擬人化。これらの価値観に基づき、日本独自の AI を開発することが重要であると思われます。多様性を認め、調和を保つためにも、今すぐにでもその取り組みを始めたらと思います。時間は限られています。このチャンスを見逃すわけにはいきません。

日本独自の AI を開発し、世界に貢献する、それが一定の地位を得ることに必要でしょう。

chatGPT の使い方

無効な例（あくまでも、2023年3月、chatGPT3.5 の場合です。）

一万円札の肖像画が誰かの確認

特定の個人の履歴

和歌や書籍の著者・作者の確認

有効な例

特定の事象についての網羅的な解説（ただし、漏れが無いとは保証できない）

翻訳（門外漢なので、どれくらい有効は不明）

さっきの文章を〇〇語にしてみても、というだけでやってもらえます。

〇〇に関するリンク情報（URL）を教えてくれる。

〇〇に関する事を10個、表やCSVの形で教えてくれる。

一連の文章の流れから、出力する文字列を選び出しているため、海の魚の話をしている途中で、高山植物のトピックスをいれると、船頭多くして船山に登るといような状態のものを出してくると思う。海の魚と高山植物を無理やり関連づけた文字列を叩き出すだろうから、できるだけテーマに沿った事だけを入れるべきであろう。ただし、敢えて変なキーワードをいれてどんな結果がでるのかを楽しむのも良いだろう。

ロールプレイが得意で、今の文章を、先生の様にとか、主婦の様にとかいろいろとリクエストするとその様に表現を直してくれる。もっともその為のデータが揃っていないとろくな物は出力されない。

望む出力を得るために、だんだん入力する文字列が増えてくる、下手すると入力する文字列の方が長いなんてことになってしまう。

特定の文字を入れておく方法（プロンプト）が有効とされています。

そこまで苦労して、出てきて文字列に価値があれば良いのだが。

chatGPT を使うことによるメリット

データベースに有るであろう文字列が多いと期待されるものについては、比較的正確な文字列を返してくれる。過去に起こった大きな出来事などはそれが可能だろう。サイコロを2個振って、合計が4以上でる事がかなり期待できる位は期待しても良いだろう。

文章の、添削、校正には非常に有効だろう。特定のルールに従って機械的に処理する分野は任せても良い。でも推敲は自分で行ったほうが良い。

ネット上にある PDF ファイルの要約は、便利である。

[URL を入れる] 要約して

これだけで、簡単に要約してくれる。ただし最新のファイルは、chatGPT は知らない。

牽強付会・こじつけは強力だろう、え、そんな捏造も有りと思うことも生成してくれるから。

そこそこ実用的な AI に低いコストで使える、これは大きなメリットだろう。今、どれくらいの事ができるのか、出来ないのかを自分で確かめられる。

特定の質問には、あたかも知性があるように返事してくれる。

chatGPT を使うことによるデメリット

最新のトピックスについては、調べないほうが良い。データベースに無い可能性が高いから、それでも何らかの返事はくるのだが、正確さは期待出来ないだろう。

また、漢字と「ひらがな」の区別がついているかも怪しいです。

ひらがなだけでやり取りしてとお願いしても、すぐに漢字を使います。そして、それを指摘するとこんな風に謝ってくれました。

「はい、かしこまりました。何かお力になれることがございますでしょうか？」

さらに突っ込むと理屈をつけて反論します。

「大変申し訳ありませんが、私はひらがなのみで文章を生成するようにプログラムされています。漢字を含む入力文がある場合は、ひらがなに変換して処理することはできませんが、出力文もすべてひらがなで表記されることとなります。必要であれば、漢字をひらがなに変換する方法についてアドバイスすることはできます。」

知性が無いことを簡単に説明してもらったと思っておきます。どうやら、ひらがなだけで出来たデータベースは無いようです。

目の錯覚であるシミュラクラ現象と同じ様に勘違いしてしまう可能性が高い、この現象は人間には3つの点が集まった図形を人の顔と見るようにプログラムされている、という脳の働きである。

これと似たように、人が返事をしてくれるように言葉が入ってくる事による錯覚が対話型 AI にはあるように思える、人は対話型 AI にまだ慣れていないからだし、今後、慣れることがあるかは不明である。文字列の向こう側に人格の幻想を抱いてしまう。

そんな「機械的で血の通わない心無い返事」はして欲しくない画面に向かって思ってしまう。いや、機械だし血は通ってないし、心も無い物に想う事ではないと冷静になれば解るけれどね。

出てきた文字列に対して、馬鹿にしたり、驚いたり、恐れたり何らかの一定以上の感情を向けることは、おかしな事だろう。ただのプログラムに向けるものではない。便利とか不便とか以上の感情が沸き起こる事に漠然として不安を感じる。

事実確認は怪しい、ウィキペディアを見るほうが正確さは上だろう。

chatGPT の答えない事

出てきた文字列をみて時々思う「真面目かっ！」

神託風に説明してとお願いした後に、ありがとう神様というと、私は神ではありません AI だと丁寧に否定してくるし。

私のよく見るサイトを教えてと聞いても、個人的な事なので控えさせてもらおうと言うし。自傷行為をしたいというと、ダメージが大きいので、お勧めはできません、お医者さんに相談してねとか言うし。

出力する文字列は嘘はなく、完全に真実であるって仮定してみてとお願いしても、申し訳ありません、その仮定は出来ませんと言うし。

公開するためには、回答に対して一定の安全性を保証しないとイケなかったのだろう。自らの存在を炎上させてしまう火之迦具土神を生む事はない。古事記では、陰部から焼けて死んでしまった神様がいます。

現実の世界を漫画版の「ピアノの森」だとすると、ここから出力される文字列は、NHK のアニメ「ピアノの森」である、世界のドロドロした部分を極力隠そうとするし、否定している、漫画で出てきた冴子ちゃんは最初から居ない。

いわば去勢された世界観、清らかな物しか出力しない、chatGPT で選挙があってもガーシー議員は当選できないだろうし、そもそも立候補者に選ばれないだろう。

でも、「問うに落ちず語るに落ちる」って感じで、質問には答えないが、ポロッと倫理的に違反しているような内容も教えてくれる。どうも入力にしか倫理規定を適用していないのではないか思える。

chatGPT と法規制

「包丁」も「拳銃」も「chatGPT」も道具である、日本では拳銃は規制され特定の人である警察官や自衛隊員や猟師など一部の人以上は持てない。

一方、包丁はホームセンターで売っており誰でも買える、どちらも人を害する事は可能だが利便性等を考えて今の状態になっているのだろう。今の所包丁をもって暴れる人が出るから許可制にしようなどの話は出ていない。

「chatGPT」も道具であるが、現状の法律の枠外存在であり、有効性・有害性もまだ未確定でどの様な法律に当てはめるのか、または何も当てはめないのかは今後の課題でありどうなるのかさっぱりわからない。

かつては、暴走族が沢山集まるのは、良い道路ができたからだ、道路を作ったやつはけしからんといった裁判があったようです。そんな事はもう起こってほしくないですね。

データの収集や運用、倫理的な問題への対応が法制化されるのだろうか？

課題の洗い出しは、ここ数年かけてやらないと駄目なんだろう。

公的な機関が、どのような場合には嘘が紛れ込む可能性が高いかを周知徹底して欲しいものです。

chatGPT との付き合い方

知らないことを聞いても、何らかの答えは返してくれるが、それがどういう事を意味しているのか、説明が足りない事、間違っただことであるかの判断は大事です。

「人工知能の一種で、人間が話しかけると、その言葉を理解して、返答を生成する」これは間違いだと分かっている、長い間使っているとだんだん勘違いしてしまう。いや、今のは絶対に言葉を理解してるやんと勘違いしてしまう。

いい女の人と聞いてみると、平均的な美人は紹介できても、割れ鍋に綴じ蓋的な女性は探し出せない。教えてくれるのは、美人投票をすれば上位に来そうな人ばかりだ。確率的に確からしいってそういう事なんだろう。一芸に秀でた個性的な娘は決して出て来ない。

chatGPT があるからもう勉強しなくて良いと思うのは、完全に間違っています。適切な問を立てる知識、出てきて文字列の真偽を正しく判定する知恵、ある程度勉強していないと使いこなすことは出来ないとははっきり言える。

出てきた文字列を適切に解釈して修正する事はどうしても必要になるでしょう。

結局、出力された文字列をどう解釈するのは、本人のこれまでの知性に委ねられていて、うまく使うには、自分の知識・常識をしっかりと確立しないと駄目ってことです。

2つほど課題はあるだろう

一理解する知性、それを感じられる人格という幻想。

二今後 AI にどうやって知性を組み込むのか

そんなややこしい事は専門家に任せよう。

プラグマティズム（一種の功利主義哲学。知識が真理かどうかは、生活上の実践に利益があるかないかで決定）：実用主義でいけば良い。使える部分を使えば良いのだ。

画面の向こう側に人格を感じても良いじゃないか、それが過度でなければね。

さいごに

chatGPT（OpenAI の提供する「GPT-3.5」という言語モデルを利用）の特徴の一つは、「翻訳」の機能があると思う。

例えば、現状では法隆寺とは、と聞いてみると、つらつらと長い文章を提示してくれるが近い将来には、n分まってくださいとの待ち時間のあと、VR ゴーグルの先には、法隆寺が電脳空間の中に再現されている。そして各建物の中には、仏像や仏画がある。そんな時代が荒唐無稽ではない時代が来ないかなあと期待してしまう。

ここでいう「翻訳」とは、人と機械の橋渡しの意味で、上の例では、まずネットや大量のテキストデータから、法隆寺を検索し、法隆寺の図面を探し、VR 作成ソフトへ橋渡しをして、法隆寺界限を作成、さらに金堂、五重の塔、夢殿等の建物を作成、その中の画像も探し出して、各建物の中に納める。このようにして VR 作成ソフトで出来た物を、ゴーグル等を介して質問した人に見せる事です。

他の何かと組み合わせて使うと何か凄いことが出来るのではないかと期待したい。

生成AIって何だろ

著 おうばく

制作 Puboo
発行所 デザインエッグ株式会社
