

目次

I 生成 AI とはどのようなものか？

- 1. AI とは何か？ 6
- 2. ChatGPT はどんなデータを学習してきたのか 8
- 3. ChatGPT 注目の新機能 10
- 4. 革新的機能「GPTs」 —自分でカスタム ChatGPT が作れる ... 12
- 5. AI が苦手なこと 14

II 読み取る AI

- 6. 文字や文章を読み取る AI 16
- 7. 音声を聴き取る AI 17
- 8. 画像の中身を認識する AI 18
- 9. リモート会議に AI 書記を導入しよう 20
- 10. 文書や動画を要約する AI 22
- 11. 長文の得意な生成 AI「Claude」を使おう 24

III 考える AI

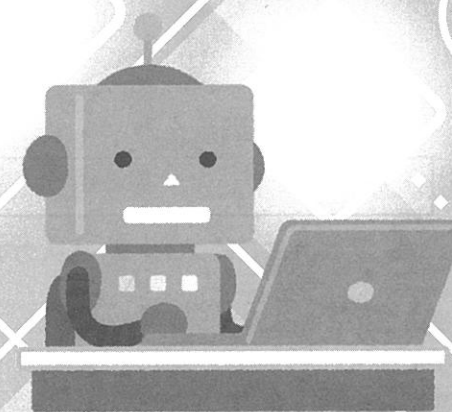
- 12. 複数に分かれて無限に議論する AI 26
- 13. コンサルティングしてくれる AI 28
- 14. アイデア出しする AI 31
- 15. AI とはどのようなスタンスでつきあうべきか 32

IV 仕事する AI

- 16. AI が言うことを聞くプロンプトのコツ 34
- 17. 企画書・リサーチ・レポートなどの書類作成する AI 36
- 18. SEO 対策つきで長文ブログ記事を書く AI 38
- 19. SNS の傾向を学んで投稿を添削してくれる AI 40

V 制作する AI

- 20. ウェブサイトを制作する AI 42
- 21. チラシ・ポスターや動画サムネールなどビジュアルを制作する AI 43
- 22. YouTube 動画などのシナリオを書く AI 44
- 23. こまかな動画編集をこなす AI 46
- 24. 企業が今すぐ始めるべき対策とは？ 48



1. AI とは何か？

「AI」「生成 AI」と簡単に言うが、なかなかつかみ所のないもので、まずどんなものかについて概念だけでも説明しておこう。

知能のベースにあるのは「認識」

AI は、以前は「人工知能」とも呼ばれており、人間の思考など知的活動を模したモノと大ざっぱに考えられている。そのベースとしてまず必要になるのは、言葉なら言葉をただのデータの羅列ではなく、意味ある塊でとらえ、その構造を理解するということだ。

我々人間が両の眼で世界を見るとき、まずは深度（距離感）つきの画像として見ているわけだが、それだけではなく、机なら机、建物なら建物、人間なら人間と意味のある塊ごとにも認識している。古い言葉で「ゲシュタルト」と言う。

我々が写真を見るとき、それは平面であるにもかかわらず、写真の中に奥行きを認識しているはずだ。これはゲシュタルトによって脳内で世界の空間を再構成していることに他ならない。漫画などより抽象的な表現であっても同じ働きを脳がしてくれるので、混乱せずに読んで作品世界に入り込めるのだ。

AI は近年この認識能力が飛躍的に上がってきた。

人間はなぜ文章が読めるのか

我々が見ているのが本である場合、「これは文字である」「文字列である」「助詞でつながって（切れて）いる単語の並びである」「文法は SOV 型で主語省略がある」「横書きだから左から右に読んでいって折り返しがある」「文脈があって、文章となっている」などの幾重もの認識の層があってその上に文章理解が成立している。

こういった言語の研究は昔からあり、「形態素分析」といったものが日本語変換ソフトに応用されてきたり、OCR といって文字の写っている画像から文字を読み取って、データとして書き出すソフトなどが開発・製品化されていた。

もとより AI 的な研究がベースにあった日本語変換はともかく、特に昔の OCR の精度はきわめて怪しいものであり、実用に耐えないという印象が強かったが、AI 技術と結びついた近年の向上の凄まじさはかなり身近に体験しているものではないだろうか。

高等教育によって AI は成長する

では、どのように向上してきたのだろうか。

まず画像の解像度自体、つまりデータ量が上がってきたという機械的なものは当然ある。それを受け入れる機械的な容量と計算能力の向上もあるということだ。

そして AI は「教育」することができる。大量のデータを与えることで、同じもの（要素）・似ているもの・全然違うものを膨大なデータに対して比較してより分け、その大量の比較の積み重ねから「認識」というものを獲得する。収集・認識・解析を繰り返すループによって自ら成長することも可能となっている（深層学習 <Deep Learning>）、人間がそのように育ってきているようにだ。基本的にデータが大量であるほど、AI は高度に成長する。ただ「喰わせる」データが良質である必要はある。

生成 AI とは何か？

さて飛躍的に向上した認識能力の上に、現在の AI は「知能」「応用」とでも言いたくなる能力が、発達してきた。これが「生成 AI」と言われる分野である。

言語の理解には幾重にもなった「多層の認識」が必要であることは述べた。現在の AI は大量のデータによる訓練によって、この認識のレイヤーをより上層へと上っていき、「理解を深めた」のである（大規模言語モデル <LLM> という）。

その結果として「人間と会話しながら、そのタスクを理解し、必要と思われる成果を出力する」といった、実に人間ライクな振る舞いまで到達し、我々の前にわかりやすいサービスとして登場したのが「ChatGPT」である。

おまけの話

ちょっとおまけの話をしたい。

岡山の駅前で噴水と一緒に自撮りして、その写真を ChatGPT に貼り付けて「この写真に写っている内容を詳しく説明してください」と聞いてみた。すると

「この写真は晴れた日に撮影された自撮り写真です。写っている男性はメガネをかけて微笑んでおり、背景には大きな噴水があります。噴水は球状の造形が特徴的で水しぶきを周囲に散らしています。噴水の背後にはビルが見え、その一つには「KITTE」と読める文字があります。男性は右手を使って自撮りをしており、画像は昼間に屋外で撮影されたと思われます」

という文章が返ってきた。きわめて詳細で、人間より優秀ではないかと思わせる。

次に、倉敷商工会議所の前で自撮り写真を撮って ChatGPT に貼り付けて、同じように



「この写真に写っている内容を説明してください」と聞いてみた。するとほとんど同じようなことを答えたのだが、唯一間違っていたのは

「背景には長岡技術科学大学と書かれた建物があります」

と言うのだ。私は「惜しいね。文字が違うよ。よく見てください」と言ったら

「申し訳ありません」「後ろに写っているものは長岡造形大学と書かれています」

と言う。長岡造形大学は私が講義している大学の名前である。人間なら「思い込みによる勘違い」で済むのだが、間違うにしても、何かちょっと弄ばれているのではないかと思うような間違いをしてきた。これはちょっと信じられないことが起きているのではないかとゾクッとしたのである。

この後、字数を指摘したり、最初の文字は「倉」だと教えたり、ChatGPT の「頑固な思い込み」を解くのに苦労した。

何がどうしてこうなったのかはわからないが、これは最初から「倉敷」だとわかっていて、なおかつ何らかの背景情報で答えを変えたようにも思える。この画像の認識、そして文字の認識を超えて、「画像の解釈」までして混乱しているようだ。

良くも悪くも AI の進歩を感じさせる出来事だった。