

ChatGPT関連情報のリサーチのコツ (15minバージョン)

2023/09/22

NOB DATA株式会社

代表取締役 大城 信晃



NOB DATA

会社紹介等

自己紹介



NOB DATA

大城信晃(twitter @doradora09)
NOB DATA株式会社 代表取締役社長 /
データサイエンティスト協会九州支部 支部長

- ・ 住まい
沖縄 -> 東京 -> 福岡
- ・ 職歴
ヤフー-> DATUM STUDIO ->
LINE Fukuoka -> NOB DATA(株) 創業
- ・ 運営コミュニティ
 - ・ Tokyo.R(2010-2016)
 - ・ fukuoka.R、PyData.Fukuoka、
オモシロAIごった煮勉強会、ChatGPT部、
他2つ



最近ではChatGPTのDX活用に興味

2023/6/14-16の3日間は幕張メッセの展示会に出展中  **NOB DATA**





NOB DATA

ChatGPT関連情報の リサーチのコツ

サマリ：ChatGPT関連情報のリサーチのコツ



NOB DATA

- **まずは本家本元のOpenAI社のChatGPT-4について調べるべし** (次点でMS AzureのOpenAI API)
 - xxGPTサービスはそもそも裏側はOpenAI社のChatGPTを使っているので、本家を把握すれば良い
- **自分で「楽しみながら」触ってみるのが習得の一番の近道**
 - 新技術は面白いと思えるかどうか、が分水嶺。または面白いと思う人を調査・普及担当にするべき
- **GPT-4の性能の外観を掴んでおく**
 - 英語でのコミュニケーションが最も得意。一説ではIQ155相当、と言われるケースもある(要出典)
- **自前でLLMを作るのは困難、基本はプロンプトで対応できる範囲にフォーカス**
 - ①オリジナルのモデル構築 -> ②ファインチューニング -> ③プロンプティング、の③を中心に調査すべし
- **発展：技術動向のウォッチ、特に大元のTransformerという技術 (GPTのT)**
 - 技術の系譜から、少し先が読め、言語モデルの技術は画像生成系の技術とも繋がっていることがわかる
 - 生成ができるということは、逆に使えば解釈にも使える(言語解釈、画像解釈)
- **大局観：政治動向やGPT-5に向けた情報収集**
 - 日本は積極活用スタンス、ヨーロッパは規制強め、アメリカはデータ取得に制限がかかりつつある
 - GPT-5は商標登録は完了、リリース時期は明かされていないが来年以降登場の可能性はある
- **最速の情報源はTwitter**
 - サムアルトマンCEO、また各種国内の情報発信アカウント(有志)のフォロー

1. まずは本家本元のOpenAI社のChatGPT-4について調べるべし NOB DATA

- xxGPTサービスはそもそも裏側はOpenAI社のChatGPTを使っているので、本家を把握すれば良い
 - なおMSはOpenAI社のChatGPT-4のクローンを保有している唯一の企業（1兆円投資している関係もあり）
-> MS copilotや github copilotシリーズの裏側もOpenAI社のChatGPTがベース。BingのAIも同じく。

本家本元のOpenAI社のChatGPT

<https://chat.openai.com/>



ChatGPTは、人々、場所、または事実について不正確な情報を提供する可能性があります。ChatGPT 8月3日バージョン

Microsoft AzureのOpenAI Service

<https://azure.microsoft.com/ja-jp/products/cognitive-services/openai-service>



2. 自分で「楽しみながら」触ってみるのが習得の一番の近道



NOB DATA

- ChatGPTを使うためには、特に専門のスキルは不要です（会話ができればOK）
- 「問う力」が求められます
- 新技術なので、自分で実際に使いながら答えを見つけるスタンスが重要です

会話を中心とするタスク

議論、作文、アイデア出し、逆質問、創作、ロールプレイ、
評価等々

プログラミングを 中心とするタスク

あらゆるプログラミング言語に精通。
計算を間違えさせたくない時はプログラムで計算させる

データ分析を 中心とするタスク

ビジネス課題を解きたい場合はデータ分析が有力候補。
ただし情報漏洩等へのケアが必要

4. 自前でLLMを作るのは困難



NOB DATA

- ChatGPT-4と同じものをゼロベースで作ろうとするとおそらく**1兆円規模の予算、7年の時間、100名以上のデータサイエンティスト(平均年俸1億円)**が必要となります(**ほぼ不可能**)
- 野良のローカルで動作するLLMもいくつか登場していますが、性能はまだChatGPT-4には遠く及びません(数年後は肉薄する可能性あり)
- ①オリジナルのモデル構築 -> ②ファインチューニング -> ③プロンプティング、の③を中心に調査すべきです
- プロンプトの例としては「**深津式**」の入力や「**グラウンディング**」という背景情報の付与がテクニックとしてあります

深津式プロンプト・システム

命令書:

あなたは**[プロの編集者]**です。

以下の制約条件と入力文をもとに**[最高の要約]**を出力してください。

制約条件:

- 文字数は**300文字程度**。
- 小学生にもわかりやすく。
- 重要なキーワードを取り残さない。
- 文章を簡潔に。

入力文:

[入力文章]

出力文:

Copy

<https://bocek.co.jp/media/exercise/chatgpt/3713/>

5. 発展：技術動向のウォッチ(1/2)

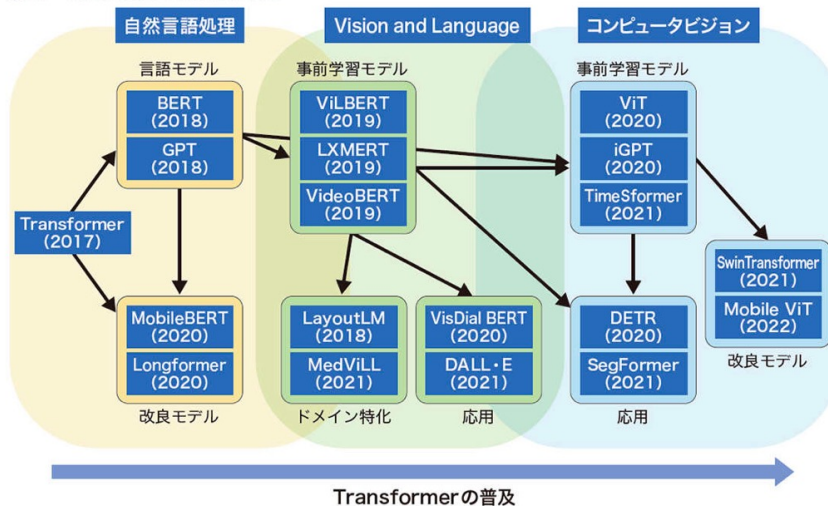


NOB DATA

- ・ 特に大元のTransformerという技術 (GPTのT)はある程度ウォッチしておいてもよいかもしれません
- ・ 下記の書籍などはおすすめです



図 1.1 : Transformer を巡る研究動向



5. 発展：技術動向のウォッチ(2/2)



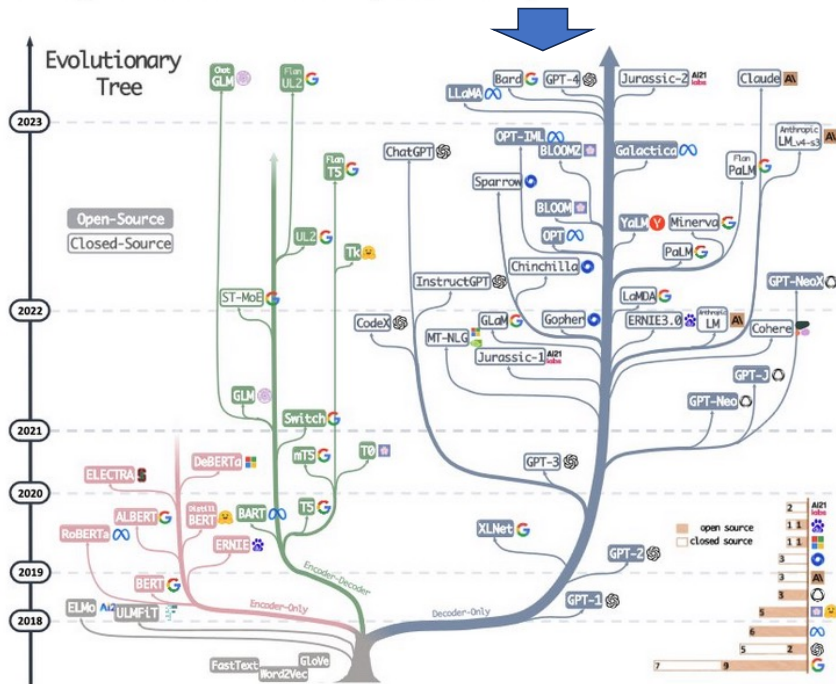
NOB DATA

- Googleもbardを作って追いかけている最中(ただし、まだまだGPT-4には敵わない)

https://twitter.com/Jeande_d/status/1673348689564254208/photo/1

Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond

3



6. 大局観：政治動向やGPT-5に向けた情報収集



NOB DATA

- 日本は比較的キャスティングボートを握りやすいポジション
 - 日本は積極活用スタンス、ヨーロッパは規制強め、アメリカはデータ取得に制限がかかりつつある
- GPT-5は商標登録は完了、リリース時期は明かされていないが来年以降登場の可能性がある
 - 11/6にOpenAI社の技術者向けのカンファレンスがある模様。必見。

TECH+ > 企業IT > 企業動向 > 開発者は200万人、Open AIが11月に初のデブコンを開催へ

開発者は200万人、Open AIが11月に初のデブコンを開催へ

掲載日 2023/09/07 12:05

著者：末岡洋子



「ChatGPT」の開発で知られるOpenAIが9月6日、初の開発者カンファレンス(デベロッパーカンファレンス)を11月6日に開催することを発表した。

<https://news.mynavi.jp/techplus/article/20230907-2766568/>

7. 最速の情報源はTwitter



NOB DATA

- 3ヶ月後に論文が出る頃にはすでに時代遅れのスピード感
- Twitterは玉石混交ですが、現時点では最速で最新情報が得られます
- サムアルトマンCEO他、日本国内の情報発信アカウント(有志)のフォローをおすすめします



<https://twitter.com/sama>



NOB DATA

ご清聴、ありがとうございました・・・！

何か個別のご相談やご質問がございましたら
info@nobdata.co.jp (NOB DATA問い合わせ窓口)
までお気軽にどうぞ。