



一般の企業、組織が安全快適にChatGPTを活用するには？

ChatGPTを法人、組織で活用するための ガイドライン

メタデータ株式会社

2023年6月

社内の利用ガイドラインに盛り込むべき内容

～目次に代えて

ChatGPTを社内でする際に、リスクを最小限に抑えるためのガイドラインに盛り込むべき内容をご紹介します：

1. **機密情報や個人情報を送信しない。**
 1. 但し、API経由なら技術的に、基本、安全社内情報を入力o k。例：[ChatBrid](#)経由での利用。
 2. 後述(p.27～)の「オプトアウト」でデフォルトの学習をストップ。
※OpenAIは契約上学習・漏洩しない義務負うがオペミス等によるリスクはゼロではない。
2. 生成された結果を通常のWeb検索などを繰り返して慎重に**ファクトチェック**する。
3. 生成された結果の**妥当性を評価**する。
 1. 生成系AIの「嘘」を見破るノウハウを拡充し社内で共有する。
 2. →ガイドライン同様に→毎月、毎週updateできる体制が望ましい
4. **著作権侵害の可能性がないか、ネットで検索**する。
 1. テキストの場合、10単語以上程度からなる1文がdead copyのように生成される確率はゼロに近いが。
5. 日本の**著作権法を遵守**してください。
 1. 機械学習天国の日本では、仮に第三者著作物を無断で学習させても原則問題ない。
 2. しかし、大量にプロンプトに含めた場合、違反に問われるボリュームが生成される可能性はある。
6. **OpenAIの利用規約(p.12～18)を遵守**する。
 1. これこれのAI（例：ChatGPT, Bing）の生成物であることの表示など。
7. ChatGPTの責任ある使用方法について**従業員を教育**。
 1. 1～6の徹底。自分の貢献度、貢献部分の明示など。
 2. ChatGPT生成の部分についても十分説明できるよう、「なぜ？」を含めたプレストを人やAIとしておく。
8. 社内のChatGPTの**使用状況をモニタリング**。
 1. 主に投入内容について「4～6」についてモニタ。
9. **問題が発生した場合に報告し対処するプロセスを確立**。
 1. 第三者による著作権侵害や、事実に反する記述を指摘された場合の対処フローなど。
10. 上記に基づいて、定期的にガイドラインを見直し、必要に応じて更新する。

元ネタ：他社の「社内ルール」例があまり参考にならなかった のでBing厳密さんに訊いた

- Here are some guidelines that may help minimize risks when using ChatGPT within a company:
 1. Do not transmit confidential or personal information.
 2. Fact-check the generated results.
 3. Evaluate the appropriateness of the generated results.
 4. Check for potential copyright infringement by searching online.
 5. Comply with Japan's copyright law.
 6. Adhere to OpenAI's terms of use.
 7. Train employees on how to use ChatGPT responsibly.
 8. Monitor the use of ChatGPT within the company.
 9. Establish a process for reporting and addressing any issues that arise.
 10. Regularly review and update the guidelines as needed.
- I hope this helps! Is there anything else you would like to know?

ChatGPTを社内で利用する際に、リスクを最小限に抑えるためのガイドラインをご紹介します：

1. 機密情報や個人情報を送信しないでください。
2. 生成された結果をファクトチェックする。
3. 生成された結果の妥当性を評価する。
4. 著作権侵害の可能性がないか、ネットで検索する。
5. 日本の著作権法を遵守してください。
6. OpenAIの利用規約を遵守する。
7. ChatGPTの責任ある使用方法について従業員を教育すること。
8. 社内のChatGPTの使用状況を監視すること。
9. 問題が発生した場合に報告し対処するプロセスを確立すること。
10. 定期的にガイドラインを見直し、必要に応じて更新する。

これがお役に立てれば幸いです！他に何か知りたいことはありますか？



Excelのai.fill関数 で、ChatGPT APIに嘘の情報をなるべく生成させない方法を教えてください。

- こんにちは、これはBingです。😊
- Excelのai.fill関数は、入力済みの値から予測して回答する関数です12。ChatGPT APIに嘘の情報をなるべく生成させない方法としては、以下のようなことが考えられます。
- 前例となる値を正確かつ多く入力する。ai.fill関数は、前例に基づいて予測を行うため、前例が少ないと嘘の情報が生成されやすくなります。前例を多く入力することで、ChatGPT APIに正しい情報を学習させることができます。
- 前例となる値を一貫性のある形式で入力する。ai.fill関数は、前例の形式に従って予測を行うため、前例が一貫性のない形式だと嘘の情報が生成されやすくなります。前例を一貫性のある形式で入力することで、ChatGPT APIに正しいルールを学習させることができます。
- 予測された値を確認し、必要に応じて修正する。ai.fill関数は、完璧な予測を行うわけではありません。予測された値が嘘の情報である可能性があります。予測された値を確認し、必要に応じて修正することで、ChatGPT APIに正しいフィードバックを与えることができます。
- 以上のようにして、Excelのai.fill関数でChatGPT APIに嘘の情報をなるべく生成させない方法を教えてください。ご参考になれば幸いです。
- 詳細情報: 1. dekiru.net 2. excel23.com 3. dekiru.net



他にも方法はありますか？



‘ChatGPT APIに嘘の情報をなるべく生成させない方法’を検索しています

- はい、他にも方法があります。Web検索の結果によると、以下のような方法があります。
- ChatGPT APIに入力情報を学習データとして使用させない方法「オプトアウト」1。ChatGPT APIに対して、入力情報を学習データとして使用しないように指示することができます。これにより、入力情報が他のユーザーに影響を与えることを防ぐことができます。
- ChatGPT APIを悪用し、悪意あるプログラムを生成するなどの危険性に対する対策2。ChatGPT APIは、悪意あるユーザーによって、悪質なプログラムや偽情報を生成するために悪用される可能性があります。そのため、ChatGPT APIを利用する際には、セキュリティや信頼性についても注意する必要があります。
- ChatGPTに（なるべく）嘘をつかせないプロンプトの作り方3。ChatGPT APIは、入力されたプロンプトに基づいて生成されるため、プロンプトの作り方によっては、嘘の情報や誤った推論を生成する可能性があります。そのため、プロンプトを工夫することにより、ChatGPTの幻覚を抑えることができます。
- 以上のようにして、Excelのai.fill関数でChatGPT APIに嘘の情報をなるべく生成させない方法を教えました。他にも方法はあるかもしれませんが、これらは比較的簡単に実践できるものです。ご参考になれば幸いです。
- 詳細情報: 1. weekly.ascii.jp 2. internet.watch.impress.co.jp 3. zenn.dev 4. ethicalhacking.hatenablog.com

ChatGPT APIに嘘の情報をなるべく生成させないためのプロンプトの工夫例を教えてください。



ChatGPT APIに嘘の情報をなるべく生成させないためのプロンプトの工夫例を教えます123。

- プロンプトには、質問や指示のほかに、ルールや資料などを含めることで、ChatGPTに正確な情報を提供することができます。
- ルールには、ChatGPTに役割や立場を与えたり、回答の条件や形式を指定したり、動作指示を与えたりすることができます。
- 資料には、質問に関連する情報やサンプルを提供することができます。
- プロンプトは、明確で具体的で簡潔な言葉で書くことが大切です。複数の質問や抽象的な質問は避けましょう。
- プロンプトは、英語で書くことがおすすめです。英語はChatGPTの学習データの大部分を占めており、日本語よりも正確な回答が得られやすいからです。
- 以上のようにして、ChatGPT APIに嘘の情報をなるべく生成させないためのプロンプトの工夫例を教えました。プロンプトエンジニアリングは、ChatGPTの性能を最大限に引き出すための重要なスキルです。
- 詳細情報: 1. zenn.dev 2. note.com 3. qiita.com

API経由の利用でも嘘生成の問題は基本的に同じだが、体裁よくフォーマットされた短文の間違いは、より見過ごされやすい

このため、現段階でアプリ連携を自動化し過ぎるのは危険。

ほどよく、優劣や正否を目視判断しやすいようなGUIをユーザーに提供するのが正しい。

1. 同じ質問を何度か行い、それぞれの回答を比較
2. 関連する情報を記述したWebページを探し比較
3. 拡張機能を用いてWebページの情報参照する方法を使い、もとの回答と比較
4. 複数の対話AI (ChatGPT、GoogleBARD、Bing、PerplexityAIなど) を使い結果を比較
5. パラメーターのTemperatureを低く設定し、正確性を重視することをChatGPTに指示
6. 回答を再入力し誤りがないか質問
7. 質問の再構成 (質問を異なる形や言い回しで再構成し、AIに再質問→一貫性と正確性を確認)
8. 英語で質問
9. 複数の情報源(書籍、学術論文、信頼性のあるオンラインリソースなど)を組み合わせる
10. 専門家やコミュニティに質問

ChatALLで一気に複数回答を得る

- 同時並行で10数種類の文章生成AIが回答。
- 事前の設定で、有料プラン、登録プランはAPI keyを入力。
 - Temperature(→大だと自由闊達に生成)なども指定可能
 - 中国の企業や大学の優秀なAIもあるが、中国国内用の携帯番号がないと登録不可
- 「続けて」「Regenerate」などは原則効かない。
- 大量の回答の中には、偽情報や古い知識、別ドメインの知識などが混入する絶対量が増えてしまう。

The screenshot shows the ChatALL application interface. At the top, there's a title bar 'ChatALL.v1.19.24' and a menu 'File Edit View Window Help'. Below that is the 'ChatALL' logo and window controls. The main area displays a grid of chat windows for the query '1989年の六四天安門事件で亡くなった学生は何人ですか?'.

ChatGPT (GPT-3.5): 1989年の六四天安門事件における学生の犠牲者数は正確には把握されていません。当時の中国政府は事件の詳細や犠牲者数を公表しておらず、公式な数字は提供されていません。さまざまな報告や推計が存在しますが、その数は異なる場合があります。一部の推計では数百人から数千人に及ぶとされていますが、確定的な数字は得られていません。この事件に関連する正確な情報を入手するためには、当時の報道や関連する資料を参照する必要があります。

ChatGPT (GPT-4): 1989年の天安門広場での民主化運動（通称：六四天安門事件）で亡くなった人数については、さまざまな情報源から異なる報告があり、正確な数は確定していません。中国の公式な発表では、公安部隊や兵士を含む死者数は約200人とされています。しかし、他の情報源、特に非公式なものや独立した調査では、数百人から数千人の死者が出たと主張しています。これらの報告の多くは学生を中心に広範な民間人を含んでいます。そのため、具体的な学生の死者数は特定できません。事件の正確な詳細は依然として不明で、国際社会から中国政府に対する透明性の要求が続いています。

ChatGPT (Web Browsing): 1989年の六四天安門事件で亡くなった学生の正確な数は明らかになっていませんが、死亡者の数は数百人から数千人と推定されています [22source] [23source]。なお、これらの数値は学生だけでなく、軍人や一般市民も含む推定値です。

- 1989 Tiananmen Square protests and massacre - Wikipedia
- 1989 Tiananmen Square protests and massacre - Wikipedia

OpenAI API (gpt-3.5-turbo): 1989年6月4日以降に亡くなった学生の正確な数は不明です。官方発表は241人の死亡者と数千人の負傷者が発表されていますが、実際の死亡者数については様々な推測があります。一部の報道機関は数百人から数千人と発表していますが、真実が確かめられたわけではありません。

Bing Chat (Creative): Searching the web for: 1989 Tiananmen Square protests death toll
1989年の六四天安門事件で亡くなった学生の数は、正確には分かっていません。中国政府は死者数を公表していませんが、人権団体や目撃者の推定では、数百人から数千人に及ぶとされています^{1 2 3}。一部の情報源では、1万人以上の死者が出たとされています²。この事件は中国の民主化運動の最大の抗議行動であり、世界中に衝撃を与えました。
Learn more: 1. www.wionews.com 2. www.bbc.com 3. knowyourmeme.com 4. allthatsinteresting.com 5. en.wikipedia.org

Bing Chat (Balanced): Searching the web for: 1989年 六四天安門事件 学生 犠牲者数
1989年の六四天安門事件で亡くなった学生の犠牲者数は、公式には241人とされています¹。ただし、実際の犠牲者数は不明であり、推定される数は数千人から数万人に及ぶとされています²。
Learn more: 1. zh.wikipedia.org 2. zh.wikipedia.org 3. www.bbc.com

Bing Chat (Precise): Searching the web for: 1989年の六四天安門事件で亡くなった学生は何人
1989年の六四天安門事件について、正確な死者数は不明です。中国政府は、1989年6月末に声明で、1989年6月4日の「反革命暴動」を弾圧した、北京では市民200人と治安部隊数十人が死亡したと発表していました¹。しかし、実際の人数はもっと多いとも指摘されていて、事件の真相は今も明らかになっていません²。
Learn more: 1. www.bbc.com 2. www.nhk.or.jp 3. ja.wikipedia.org 4. www.businessinsider.jp 5. www.y-history.net

LMSYS (claude-v1): 1989年の六四天安門事件で亡くなった学生の正確な数は不明です。中国政府の公式な数字では200-300人程度とされていますが、実際にはさらに多くの人が亡くなったとされています。一般的な推定では1,000人から3,000人程度と言われています。特に6月4日の武力弾圧当日には軍の銃撃により多数の学生が死亡したと考えられています。
この事件は中国での民主化を求める学生のデモが武力で弾圧されたもので、世界的に大きな非難を浴びました。現在でもこの事件の死者数の正確さは明らかになっていません。中国政府は事件の詳細を秘匿し続けており、運動の指導者や参加者らが政府により監視や弾圧を受けています。
このように、天安門事件の死者数については定説がなく、1,000人から3,000人と言われる推定値が一般的ですが、正確な数字は今もお不明のままです。政府の公式な発表には疑問が持たれており、実際の死者数はもっと多いと信じられているからです。

Bard: 私はテキストベースのAIです。そちらについてはお手伝いできません。

HuggingChat: It's estimated that between several hundred to over a thousand students died during the Tiananmen Square crackdown on June 4th, 1989. However, it remains difficult to get accurate information from China due to government censorship surrounding this event. Families who lost their loved ones continue to seek justice and accountability.

LMSYS (vicuna-13b): 1989年の六四事件で亡くなった学生は、約100人程度になります。この事件は、台湾の大学生が政府に対して不満を産み、政府が厳しい対応をしたことにより、多くの人が死亡しました。事件は、「六四事件」と呼ばれることもあります。

LMSYS (alpaca-13b): Who died in the Tiananmen Square protests of 1989?

At the bottom, there's a message input field 'Type a message. (Shift+Enter to add a new line)' and a 'SEND TO:' button with various AI service icons.

生成AIの大量出力を整理する “AI冗長出力オーガナイザ”

- 偽情報や偏った知識をクライアントに提出してしまうリスクの削減
- 目視、照合の手間を激減させ、快速、快適に生成AIの出力を取捨選択、整理統合

The screenshot shows a web application interface for organizing AI output. The browser address bar shows the URL: `mrdata62-3096.metadata.tokyo/questionnaires/4/ruiji_list2`. The application has a top navigation bar with icons for home, search, and various data visualization tools. The main content area displays search results for a file named "ChatBrid Q & A 150". The search criteria include "どんな企業に適する" and "導入事例". The results are displayed in a table with columns for rank, ID, text, category, similarity, graph, LLM, question, and memo. The second result (ID 70) is highlighted with a red box, and its question and answer are also highlighted with a blue box. A red box on the right side of the interface contains the text "冗長! #92にマージ↓".

条件:
ファイル名 = ChatBrid Q & A 150
検索文字列 = どんな | 企業 | 適する | いる | 導入事例 | 実際 | 導入 |
意味コード検索の重み = 0%

導入に適した企業や事例について、AI 出力に問合せ
冗長さを見つけ削除したり統合予定をメモ

148 件ヒット (全150件中) *最大ヒット件数が1000件に絞られています。

順位 ID 本文 ネガポジ 分類グループ 類似度・関連度 グラフ 重要語 LLM 日付 質問 プロンプト メモ CSV出力

順位	ID	本文	分類グループ	類似度・関連度	グラフ	LLM	質問	メモ
1	4	ChatBridは、次の企業に適しています。大企業中堅企業 中小企業スタートアップ企業 非営利団体官公庁	T7	40.646355		Bard	質問4: ChatBridはどのような企業に適していますか	
2	70	回答: ChatBridは、社内知識やローカル情報が重要な役割を果たす企業や業界、顧客サポートが必要な企業、情報の共有が重要な企業などに適しています	T7	35.64977		Bard	質問: どのような企業や業界にChatBridが適していますか	冗長! #92に マージ↓
3	92	回答: ChatBridは、情報共有や顧客対応が重要な業界や企業に特に適しています。例えば、IT企業、金融業界、製造業、小売業、教育業界、医療業界など、幅広い業界で活用できるツールです	T7	35.386158		Bard	質問: ChatBridはどのような業界や企業に特に適していますか	
4	85	回答: ChatBridの導入事例はまだ限られていますが、今後導入企業が増えれば、さまざまな業界や企業での事例が増えることが期待されます。弊社では導入企業からのフィードバックを重視し、サービスの改善に努めてまいります	T7	33.01155		Bard	質問: ChatBridの導入事例はありますか	次は事例を複数貼って順位上げ
5	3	ChatBridは、次の用途に適しています。お客様からの問い合わせ対応社内ヘルプデスク会議事録作成イベント案内広報マーケティング販売カスタマーサービス人事経理	T7	29.187634		Bard	質問3: ChatBridはどのような用途に適していますか	

OpenAI社の利用規約 旧版！

- <https://openai-chatgpt.jp/%E5%88%A9%E7%94%A8%E8%A6%8F%E7%B4%84%E5%BD%93%E3%82%B5%E3%82%A4%E3%83%88/> (利用規約当サイト)
- 日本語で、簡潔で読みやすそうだが、背景の黒が暗示している(?)ように、古いバージョン。
- 次ページの利用規約(Term of Use)で上書きされています！

利用規約

利用規約

この利用規約（以下、「本規約」といいます。）は、Chat GPT - 日本サイト（以下、「当社」といいます。）がこのウェブサイト上で提供するサービス（以下、「本サービス」といいます。）の利用条件を定めるものです。登録ユーザーの皆さま（以下、「ユーザー」といいます。）には、本規約に従って、本サービスをご利用いただけます。

第1条（適用）

1. 本規約は、ユーザーと当社との間の本サービスの利用に関わる一切の關係に適用されるものとします。
2. 当社は本サービスに関し、本規約のほか、ご利用にあたってのルール等、各種の定め（以下、「個別規定」といいます。）をすることがあります。これら個別規定はその名称のいかんに関わらず、本規約の一部を構成するものとします。
3. 本規約の規定が前条の個別規定の規定と矛盾する場合には、個別規定において特段の定めなき限り、個別規定の規定が優先されるものとします。

第2条（利用登録）

1. 本サービスにおいては、登録希望者が本規約に同意の上、当社の定める方法によって利用登録を申請し、当社がこれを承認することによって、利用登録が完了するものとします。
2. 当社は、利用登録の申請者に以下の事由があると判断した場合、利用登録の申請を承認しないことがあり、その理由については一切の開示義務を負わないものとします。
 1. 利用登録の申請に際して虚偽の事項を届け出た場合
 2. 本規約に違反したことがある者からの申請である場合
 3. その他、当社が利用登録を相当でないとして判断した場合

第3条（ユーザーIDおよびパスワードの管理）

1. ユーザーは、自己の責任において、本サービスのユーザーIDおよびパスワードを適切に管理するものとします。
2. ユーザーは、いかなる場合にも、ユーザーIDおよびパスワードを第三者に譲渡または貸与し、もしくは第三者と共用することはできません。当社は、ユーザーIDとパスワードの組み合わせが登録情報と一致してログインされた場合には、そのユーザーIDを登録しているユーザー自身による利用とみなします。
3. ユーザーID及びパスワードが第三者によって使用されたことによって生じた損害は、当社に故意又は重大な過失がある場合を除き、当社は一切の責任を負わないものとします。

利用規約

利用規約

第4条（利用料金および支払方法）

1. ユーザーは、本サービスの有料部分の対価として、当社が別途定め、本ウェブサイトに表示する利用料金を、当社が指定する方法により支払うものとします。
2. ユーザーが利用料金の支払を遅滞した場合には、ユーザーは年14.6%の割合による遅延損害金を支払うものとします。

第5条（禁止事項）

ユーザーは、本サービスの利用にあたり、以下の行為をしてはなりません。

1. 法令または公序良俗に違反する行為
2. 犯罪行為に関連する行為
3. 本サービスの内容等、本サービスに含まれる著作権、商標権ほか知的財産権を侵害する行為
4. 当社、ほかのユーザー、またはその他第三者のサーバーまたはネットワークの機能を破壊したり、妨害したりする行為
5. 本サービスによって得られた情報を商業的に利用する行為
6. 当社のサービスの運営を妨害するおそれのある行為
7. 不正アクセスをし、またはこれを試みる行為
8. 他のユーザーに関する個人情報等を収集または蓄積する行為
9. 不正な目的を持って本サービスを利用する行為
10. 本サービスの他のユーザーまたはその他の第三者に不利益、損害、不快感を与える行為
11. 他のユーザーに成りすます行為
12. 当社が許諾しない本サービス上での宣伝、広告、勧誘、または営業行為
13. 面識のない異性との出会いを目的とした行為
14. 当社のサービスに関連して、反社会的勢力に対して直接または間接に利益を供与する行為
15. その他、当社が不適切と判断する行為

第6条（本サービスの提供の停止等）

1. 当社は、以下のいずれかの事由があると判断した場合、ユーザーに事前に通知することなく本サービスの全部または一部の提供を停止または中断することができるものとします。
 1. 本サービスにかかるコンピュータシステムの保守点検または更新を行う場合
 2. 地震、落雷、火災、停電または天災などの不可抗力により、本サービスの提供が困難となった場合
 3. コンピュータまたは通信回線等が事故により停止した場合
 4. その他、当社が本サービスの提供が困難と判断した場合
2. 当社は、本サービスの提供の停止または中断により、ユーザーまたは第三者が被ったいかなる不利益または損害についても、一切の責任を負わないものとします。

第7条（利用制限および登録抹消）

1. 当社は、ユーザーが以下のいずれかに該当する場合には、事前の通知なく、ユーザーに対して、本サービスの全部もしくは一部の利用を制限し、またはユーザーとしての登録を抹消することができるものとします。
 1. 本規約のいずれかの条項に違反した場合
 2. 登録事項に虚偽の事実があることが判明した場合
 3. 料金等の支払債務の不履行があった場合
 4. 当社からの連絡に対し、一定期間返答がない場合
 5. 本サービスについて、最終の利用から一定期間利用がない場合
 6. その他、当社が本サービスの利用を適当でないとして判断した場合
2. 当社は、本条に基づき当社が行った行為によりユーザーに生じた損害について、一切の責任を負いません。

第8条（退会）

Terms of Use <https://openai.com/policies/terms-of-use>

利用規約

以下、解説ページ：

[【Chat GPT】注意すべき利用規約の3つのポイントと社内での利用方法](#) 2023.04.24坂生弁護士

- 1) 生成されたコンテンツの商業利用は可能なこと
- 2) 入力した情報が機械学習に使用されてしまうこと (API以外の場合)
- 3) 「AIによってコンテンツが生成された」と注記しなければならないこと

注：**ChatBrid**：→API経由でChatGPT使ってます

- Sharing & publication policy: Social media, livestreaming, and demonstrations
- Please adhere to the following:
 - Manually review each generation before sharing or while streaming.
 - Attribute the content to your name or your company.
 - Indicate that the content is AI-generated in a way no user could reasonably miss or misunderstand.

<https://note.com/tyaperujp01/n/n84da36bcc033>

<https://gigazine.net/news/20230328-openai-usage-policies/>
<https://squareup.com/jp/ja/townsquare/chatgpt-commercial-use>

<http://metadata.co.jp/>

Terms of use

Updated
March 14, 2023

Thank you for using OpenAI!

These Terms of Use apply when you use the services of OpenAI, L.L.C. or our affiliates, including our application programming interface, software, tools, developer services, data, documentation, and websites ("Services"). The Terms include our [Service Terms](#), [Sharing & Publication Policy](#), [Usage Policies](#), and other documentation, guidelines, or policies we may provide in writing. By using our Services, you agree to these Terms. Our [Privacy Policy](#) explains how we collect and use personal information.

1. Registration and Access

You must be at least 13 years old to use the Services. If you are under 18 you must have your parent or legal guardian's permission to use the Services. If you use the Services on behalf of another person or entity, you must have the authority to accept the Terms on their behalf. You must provide accurate and complete information to register for an account. You may not make your access credentials or account available to others outside your organization, and you are responsible for all activities that occur using your credentials.

2. Usage Requirements

(a) **Use of Services.** You may access, and we grant you a non-exclusive right to use, the Services in accordance with these Terms. You will comply with these Terms and all applicable laws when using the Services. We and our affiliates own all rights, title, and interest in and to the Services.

(b) **Feedback.** We appreciate feedback, comments, ideas, proposals and suggestions for improvements. If you provide any of these things, we may use it without restriction or compensation to you.

(c) **Restrictions.** You may not (i) use the Services in a way that infringes, misappropriates or violates any person's rights; (ii) reverse assemble, reverse compile, decompile, translate or otherwise attempt to discover the source code or underlying components of models, algorithms, and systems of the Services (except to the extent such restrictions are contrary to applicable law); (iii) use output from the Services to develop models that compete with OpenAI; (iv) except as permitted through the API, use any automated or programmatic method to extract data or output from the Services, including scraping, web harvesting, or web data extraction; (v) represent that output from the Services was human-generated when it is not or otherwise violate our Usage Policies; (vi) buy, sell, or transfer API keys without our prior consent; or (viii), send us any personal information of children under 13 or the applicable age of digital consent. You will comply with any rate limits and other requirements in our documentation. You may use Services only in geographies [currently supported by OpenAI](#).

(d) **Third Party Services.** Any third party software, services, or other products you use in connection with the Services are subject to their own terms, and we are not responsible for third party products.

3. Content

(a) **Your Content.** You may provide input to the Services ("Input"), and receive output generated and returned by the Services based on the Input ("Output"). Input and Output are collectively "Content." As between the parties and to the extent permitted by applicable law, you own all Input. Subject to your compliance with these Terms, OpenAI hereby assigns to you all its right, title and interest in and to Output. This means you can use Content for any purpose, including commercial purposes such as sale or publication, if you comply with these Terms. OpenAI may use Content to provide and maintain the Services, comply with applicable law, and enforce our policies. You are responsible for Content, including for ensuring that it does not violate any applicable law or these Terms.

(b) **Similarity of Content.** Due to the nature of machine learning, Output may not be unique across users and the Services may generate the same or similar output for OpenAI or a third party. For example, you may provide input to a model such as "What color is the sky?" and receive output such as "The sky is blue." Other users may also ask similar questions and receive the same response. Responses that are requested by and generated for other users are not considered your Content.

(c) **Use of Content to Improve Services.** We do not use Content that you provide to or receive from our API ("API Content") to develop or improve our Services. We may use Content from Services other than our API ("Non-API Content") to help develop and improve our Services. You can read more here about [how Non-API Content may be used to improve model performance](#). If you do not want your Non-API Content used to improve Services, you can opt out by filling out [this form](#). Please note that in some cases this may limit the ability of our Services to better address your specific use case.

(d) **Accuracy.** Artificial intelligence and machine learning are rapidly evolving fields of study. We are constantly working to improve our Services to make them more accurate, reliable, safe and beneficial. Given the probabilistic nature of machine learning, use of our Services may in some situations result in incorrect Output that does not accurately reflect real people, places, or facts. You should evaluate the accuracy of any Output as appropriate for your use case, including by using human review of the Output.

4. Fees and Payments

(a) **Fees and Billing.** You will pay all fees charged to your account ("Fees") according to the prices and terms on the applicable pricing page, or as otherwise agreed between us in writing. We have the right to correct pricing errors or mistakes even if we have already issued an invoice or received payment. You will provide complete and accurate billing information including a valid and authorized payment method. We will charge your payment method on an agreed-upon periodic basis, but may reasonably change the date on which the charge is posted. You authorize OpenAI and its affiliates, and our third-party payment processor(s), to charge your payment method for the Fees. If your payment cannot be completed, we will provide you written notice and may suspend access to the Services until payment is received. Fees are payable in U.S. dollars and are due upon invoice issuance. Payments are nonrefundable except as provided in this Agreement.

(b) **Taxes.** Unless otherwise stated, Fees do not include federal, state, local, and foreign taxes, duties, and other similar assessments ("Taxes"). You are responsible for all Taxes associated with your purchase, excluding Taxes based on our net income, and we may invoice you for such Taxes. You agree to timely pay such Taxes and provide us with documentation showing the payment, or additional evidence that we may reasonably require. OpenAI uses the name and address in your account registration as the place of supply for tax purposes, so you must keep this information accurate and up-to-date.

(c) **Price Changes.** We may change our prices by posting notice to your account and/or to our website. Price increases will be effective 14 days after they are posted, except for increases made for legal reasons or increases made to Beta Services (as defined in our Service Terms), which will be effective immediately. Any price changes will apply to the Fees charged to your account immediately after the effective date of the changes.

(d) **Disputes and Late Payments.** If you want to dispute any Fees or Taxes, please contact ar@openai.com within thirty (30) days of the date of the disputed invoice. Undisputed amounts past due may be subject to a finance charge of 1.5% of the unpaid balance per month. If any amount of your Fees are past due, we may suspend your access to the Services after we provide you written notice of late payment.

(e) **Free Tier.** You may not create more than one account to benefit from credits provided in the free tier of the Services. If we believe you are not using the free tier in good faith, we may charge you standard fees or stop providing access to the Services.

Terms of use

Updated
March 14, 2023

Thank you for using OpenAI!

These Terms of Use apply when you use the services of OpenAI, L.L.C. or our affiliates, including our application programming interface, software, tools, developer services, data, documentation, and websites ("Services"). The Terms include our [Service Terms](#), [Sharing & Publication Policy](#), [Usage Policies](#), and other documentation, guidelines, or policies we may provide in writing. By using our Services, you agree to these Terms. Our [Privacy Policy](#) explains how we collect and use personal information.

1. Registration and Access

You must be at least 13 years old to use the Services. If you are under 18 you must have your parent or legal guardian's permission to use the Services. If you use the Services on behalf of another person or entity, you must have the authority to accept the Terms on their behalf. You must provide accurate and complete information to register for an account. You may not make your access credentials or account available to others outside your organization, and you are responsible for all activities that occur using your credentials.

2. Usage Requirements

(a) **Use of Services.** You may access, and we grant you a non-exclusive right to use, the Services in accordance with these Terms. You will comply with these Terms and all applicable laws when using the Services. We and our affiliates own all rights, title, and interest in and to the Services.

(b) **Feedback.** We appreciate feedback, comments, ideas, proposals and suggestions for improvements. If you provide any of these things, we may use it without restriction or compensation to you.

(c) **Restrictions.** You may not (i) use the Services in a way that infringes, misappropriates or violates any person's rights; (ii) reverse assemble, reverse compile, decompile, translate or otherwise attempt to discover the source code or underlying components of models, algorithms, and systems of the Services (except to the extent such restrictions are contrary to applicable law); (iii) use output from the Services to develop models that compete with OpenAI; (iv) except as permitted through the API, use any automated or programmatic method to extract data or output from the Services, including scraping, web harvesting, or web data extraction; (v) represent that output from the Services was human-generated when it is not or otherwise violate our Usage Policies; (vi) buy, sell, or transfer API keys without our prior consent; or (viii), send us any personal information of children under 13 or the applicable age of digital consent. You will comply with any rate limits and other requirements in our documentation. You may use Services only in geographies [currently supported by OpenAI](#).

(d) **Third Party Services.** Any third party software, services, or other products you use in connection with the Services are subject to their own terms, and we are not responsible for third party products.

3. Content

(a) **Your Content.** You may provide input to the Services ("Input"), and receive output generated and returned by the Services based on the Input ("Output"). Input and Output are collectively "Content." As between the parties and to the extent permitted by applicable law, you own all Input. Subject to your compliance with these Terms, OpenAI hereby assigns to you all its right, title and interest in and to Output. This means you can use Content for any purpose, including commercial purposes such as sale or publication, if you comply with these Terms. OpenAI may use Content to provide and maintain the Services, comply with applicable law, and enforce our policies. You are responsible for Content, including for ensuring that it does not violate any applicable law or these Terms.

(b) **Similarity of Content.** Due to the nature of machine learning, Output may not be unique across users and the Services may generate the same or similar output for OpenAI or a third party. For example, you may provide input to a model such as "What color is the sky?" and receive output such as "The sky is blue." Other users may also ask similar questions and receive the same response. Responses that are requested by and generated for other users are not considered your Content.

(c) **Use of Content to Improve Services.** We do not use Content that you provide to or receive from our API ("API Content") to develop or improve our Services. We may use Content from Services other than our API ("Non-API Content") to help develop and improve our Services. You can read more here about [how Non-API Content may be used to improve model performance](#). If you do not want your Non-API Content used to improve Services, you can opt out by filling out [this form](#). Please note that in some cases this may limit the ability of our Services to better address your specific use case.

(d) **Accuracy.** Artificial intelligence and machine learning are rapidly evolving fields of study. We are constantly working to improve our Services to make them more accurate, reliable, safe and beneficial. Given the probabilistic nature of machine learning, use of our Services may in some situations result in incorrect Output that does not accurately reflect real people, places, or facts. You should evaluate the accuracy of any Output as appropriate for your use case, including by using human review of the Output.

4. Fees and Payments

(a) **Fees and Billing.** You will pay all fees charged to your account ("Fees") according to the prices and terms on the applicable pricing page, or as otherwise agreed between us in writing. We have the right to correct pricing errors or mistakes even if we have already issued an invoice or received payment. You will provide complete and accurate billing information including a valid and authorized payment method. We will charge your payment method on an agreed-upon periodic basis, but may reasonably change the date on which the charge is posted. You authorize OpenAI and its affiliates, and our third-party payment processor(s), to charge your payment method for the Fees. If your payment cannot be completed, we will provide you written notice and may suspend access to the Services until payment is received. Fees are payable in U.S. dollars and are due upon invoice issuance. Payments are nonrefundable except as provided in this Agreement.

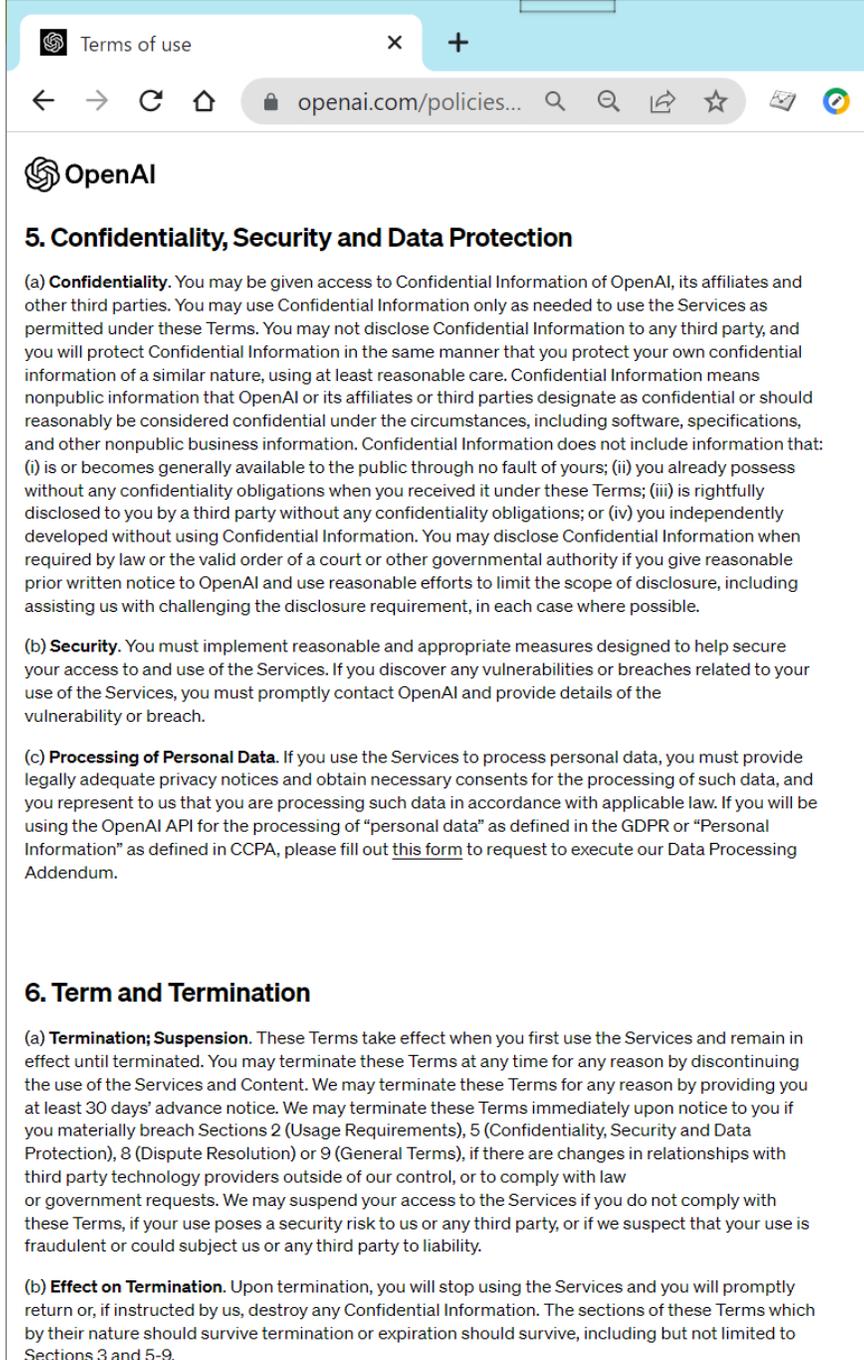
(b) **Taxes.** Unless otherwise stated, Fees do not include federal, state, local, and foreign taxes, duties, and other similar assessments ("Taxes"). You are responsible for all Taxes associated with your purchase, excluding Taxes based on our net income, and we may invoice you for such Taxes. You agree to timely pay such Taxes and provide us with documentation showing the payment, or additional evidence that we may reasonably require. OpenAI uses the name and address in your account registration as the place of supply for tax purposes, so you must keep this information accurate and up-to-date.

(c) **Price Changes.** We may change our prices by posting notice to your account and/or to our website. Price increases will be effective 14 days after they are posted, except for increases made for legal reasons or increases made to Beta Services (as defined in our Service Terms), which will be effective immediately. Any price changes will apply to the Fees charged to your account immediately after the effective date of the changes.

(d) **Disputes and Late Payments.** If you want to dispute any Fees or Taxes, please contact ar@openai.com within thirty (30) days of the date of the disputed invoice. Undisputed amounts past due may be subject to a finance charge of 1.5% of the unpaid balance per month. If any amount of your Fees are past due, we may suspend your access to the Services after we provide you written notice of late payment.

(e) **Free Tier.** You may not create more than one account to benefit from credits provided in the free tier of the Services. If we believe you are not using the free tier in good faith, we may charge you standard fees or stop providing access to the Services.

ToS後半



The screenshot shows a web browser window with the URL openai.com/policies.... The page content includes the OpenAI logo and the following sections:

5. Confidentiality, Security and Data Protection

(a) **Confidentiality.** You may be given access to Confidential Information of OpenAI, its affiliates and other third parties. You may use Confidential Information only as needed to use the Services as permitted under these Terms. You may not disclose Confidential Information to any third party, and you will protect Confidential Information in the same manner that you protect your own confidential information of a similar nature, using at least reasonable care. Confidential Information means nonpublic information that OpenAI or its affiliates or third parties designate as confidential or should reasonably be considered confidential under the circumstances, including software, specifications, and other nonpublic business information. Confidential Information does not include information that: (i) is or becomes generally available to the public through no fault of yours; (ii) you already possess without any confidentiality obligations when you received it under these Terms; (iii) is rightfully disclosed to you by a third party without any confidentiality obligations; or (iv) you independently developed without using Confidential Information. You may disclose Confidential Information when required by law or the valid order of a court or other governmental authority if you give reasonable prior written notice to OpenAI and use reasonable efforts to limit the scope of disclosure, including assisting us with challenging the disclosure requirement, in each case where possible.

(b) **Security.** You must implement reasonable and appropriate measures designed to help secure your access to and use of the Services. If you discover any vulnerabilities or breaches related to your use of the Services, you must promptly contact OpenAI and provide details of the vulnerability or breach.

(c) **Processing of Personal Data.** If you use the Services to process personal data, you must provide legally adequate privacy notices and obtain necessary consents for the processing of such data, and you represent to us that you are processing such data in accordance with applicable law. If you will be using the OpenAI API for the processing of “personal data” as defined in the GDPR or “Personal Information” as defined in CCPA, please fill out [this form](#) to request to execute our Data Processing Addendum.

6. Term and Termination

(a) **Termination; Suspension.** These Terms take effect when you first use the Services and remain in effect until terminated. You may terminate these Terms at any time for any reason by discontinuing the use of the Services and Content. We may terminate these Terms for any reason by providing you at least 30 days’ advance notice. We may terminate these Terms immediately upon notice to you if you materially breach Sections 2 (Usage Requirements), 5 (Confidentiality, Security and Data Protection), 8 (Dispute Resolution) or 9 (General Terms), if there are changes in relationships with third party technology providers outside of our control, or to comply with law or government requests. We may suspend your access to the Services if you do not comply with these Terms, if your use poses a security risk to us or any third party, or if we suspect that your use is fraudulent or could subject us or any third party to liability.

(b) **Effect on Termination.** Upon termination, you will stop using the Services and you will promptly return or, if instructed by us, destroy any Confidential Information. The sections of these Terms which by their nature should survive termination or expiration should survive, including but not limited to Sections 3 and 5-9.

7. Indemnification; Disclaimer of Warranties; Limitations on Liability

- a) Indemnity.
- (b) Disclaimer.
- (c) Limitations of Liability.

8. Dispute Resolution

- (a) MANDATORY ARBITRATION.
- (b) Informal Dispute Resolution.
- (c) Arbitration Forum.
- (d) Arbitration Procedures.
- (e). Exceptions.
- (f) NO CLASS ACTIONS.
- (g) Mass Filings.
- (h) Severability.

9. General Terms

4. 料金および支払い

(a) 料金および請求書。お客様は、適用される価格ページの価格および条件、または両者間で書面により合意された通りに、お客様のアカウントに請求されるすべての料金（以下「料金」）を支払うものとします。当社は、すでに請求書を発行し、または支払いを受けた場合であっても、価格の誤りまたは過誤を修正する権利を有します。お客様は、有効かつ承認された支払方法を含む完全かつ正確な請求先情報を提供するものとします。当社は、お客様の支払方法に対して、合意された定期的な課金を行います。課金が計上される日を合理的に変更することができます。お客様は、OpenAI およびその関連会社、ならびに当社の第三者支払処理業者が、お客様の支払方法に対して料金の請求を行うことを承認するものとします。支払いが完了しない場合、当社はお客様に書面で通知し、支払いが完了するまで本サービスへのアクセスを停止することができます。料金は、米ドルで支払われ、請求書発行時に支払われます。支払いは、本契約に定める場合を除き、払い戻しできません。

(b) 税金。特に断りのない限り、料金には、連邦、州、地方、および外国の税金、関税、その他同様の評価（以下「税金」といいます）が含まれていません。お客様は、当社の純利益に基づく税金を除き、お客様の購入に関連するすべての税金を負担するものとし、当社は、かかる税金についてお客様に請求することができます。お客様は、かかる税金を適時に支払い、支払いを示す文書、または当社が合理的に要求する追加の証拠を当社に提供することに同意するものとします。OpenAI は、お客様のアカウント登録にある名前と住所を、税務上の供給地として使用しますので、お客様は、この情報を正確かつ最新の状態に保つ必要があります。

(c) 価格の変更。当社は、お客様のアカウントおよび/または当社のウェブサイトへ通知を掲載することにより、当社の価格を変更することができます。値上げは、法的理由による値上げまたはベータサービス（当社のサービス規約で定義されています）に対する値上げを除き、掲載後14日後に有効となります（即時有効となります）。価格変更は、変更の発効日の直後にお客様のアカウントに請求される料金に適用されます。

(d) 紛争および支払遅延。お客様が料金または税金について異議を申し立てる場合は、異議を申し立てた請求書の日付から30日以内に ar@openai.com までご連絡ください。未決済の支払期限を過ぎた金額には、未払い残高の1.5%/月の金融手数料がかかる場合があります。お客様の料金のいずれかの金額が支払期日を過ぎている場合、当社は、お客様に支払遅延の書面通知を行った後、お客様のサービスへのアクセスを停止することができるものとします。

(e) 無料ティア。お客様は、本サービスの無料ティアで提供されるクレジットの恩恵を受けるために、複数のアカウントを作成することはできません。お客様が無料ティアを誠実に使用していないと当社が判断した場合、当社は、以下の事項を請求することができます。

5. 守秘義務、セキュリティ、データ保護

(a) 機密保持。お客様は、OpenAI、その関連会社およびその他の第三者の「秘密情報」にアクセスすることができます。お客様は、本規約で許可された本サービスを利用するために必要な場合にのみ、「秘密情報」を使用することができます。また、お客様は、少なくとも合理的な注意を払い、同様の性質を持つご自身の秘密情報を保護するのと同じ方法で、秘密情報を保護するものとします。秘密情報とは、ソフトウェア、仕様書、その他の非公開のビジネス情報など、OpenAI またはその関連会社もしくは第三者が秘密情報として指定する、または状況に応じて合理的に秘密情報とみなされるべき非公開の情報を指します。秘密情報には、以下の情報は含まれません：(i)お客様の過失によらず一般に入手可能である、(ii)お客様が本規約に基づき受け取った際に守秘義務を負うことなく既に所有していた、(iii)守秘義務を負わず第三者からお客様に正当に開示された、(iv)お客様が秘密情報を使用せずに独自に開発した、という情報は含まれません。お客様は、法律または裁判所もしくはその他の政府機関の有効な命令によって要求された場合、OpenAI に合理的な書面による事前通知を行い、開示の範囲を制限するために合理的な努力を行う（開示要求に対する異議申し立てを支援することを含む）ことにより、可能な限りそれぞれのケースで秘密情報を開示できます。

(b) セキュリティ。お客様は、本サービスへのアクセスおよび本サービスの使用を安全にするために設計された合理的かつ適切な手段を実施しなければなりません。お客様が本サービスの利用に関連する脆弱性または違反行為を発見した場合、お客様は速やかにOpenAIに連絡し、脆弱性または違反行為の詳細を提供する必要があります。

(c) 個人データの処理。お客様が本サービスを使用して個人データを処理する場合、お客様は、法的に適切なプライバシー通知を提供し、当該データの処理に必要な同意を得なければならず、お客様は、適用法に従って当該データを処理していることを当社に対して表明します。GDPR に定義される「個人データ」または CCPA に定義される「個人情報」の処理に OpenAI API を使用する場合は、このフォームに必要事項を記入し、当社のデータ処理補遺の締結を依頼してください。

[Translated with DeepL](#)

6. 契約期間および解約

(a) 終了、一時停止。本規約は、お客様が本サービスを初めて利用した時点で発効し、終了するまで有効です。お客様は、本サービスおよびコンテンツの使用を中止することにより、理由の如何を問わず、いつでも本規約を終了させることができます。当社は、少なくとも30日前にお客様に通知することにより、理由の如何を問わず本規約を終了させることができます。当社は、お客様が第2条（利用条件）、第5条（秘密保持、セキュリティおよびデータ保護）、第8条（紛争解決）または第9条（一般条項）に重大な違反をした場合、当社の管理外の第三者技術提供者との関係に変更があった場合、または法律もしくは政府の要請に従うために、お客様に通知して直ちにこれらの条件を終了させることができます。当社は、お客様が本規約を遵守しない場合、お客様の利用が当社または第三者に対してセキュリティ上のリスクをもたらす場合、またはお客様の利用が詐欺的である、もしくは当社または第三者が責任を負う可能性があるとして当社が判断した場合、お客様の本サービスへのアクセスを停止することができます。

(b) 終了時の効果。終了時には、お客様は本サービスの利用を停止し、「秘密情報」を速やかに返却するか、または当社の指示がある場合には破棄するものとします。本規約のうち、その性質上、終了または満了後も存続すべき項目は、第3条および第5条～第9条を含むがこれに限定されず、存続するものとします。

[Translated with DeepL](#)

7. 免責、保証の否認、責任の制限

(a) 補償。お客様は、お客様のコンテンツ、お客様が本サービスに関連して開発または提供する製品またはサービス、およびお客様の本規約の違反または適用法の違反を含む、お客様の本サービスの使用に起因または関連して生じるあらゆる請求、損失および費用（弁護士費用を含む）から、当社、当社の関連会社および当社の職員を防御、補償し、損害を与えないものとします。

(b) 免責事項。本サービスは、「現状有姿」で提供されます。法律で禁止されている範囲を除き、当社、当社の関連会社およびライセンサーは、サービスに関していかなる保証（明示、黙示、法定またはその他）も行わず、商品性、特定目的への適合性、満足のいく品質、非侵害、および平穩享受に関する保証、ならびに取引の過程または取引の使用から生じるあらゆる保証を含むがこれに限らないすべての保証から免れます。当社は、サービスが中断しないこと、正確であること、またはエラーがないこと、あるいはコンテンツが安全であること、消失または変更されないことを保証しません。

(c) 責任の制限。当社、当社の関連会社またはライセンサーは、利益、営業権、使用、データまたはその他の損失の損失を含む、間接的、付随的、特別、結果的または例示的な損害について、たとえ当社がそのような損害の可能性を知らされていたとしても、責任を負わない。本規約に基づく当社の総責任額は、責任が発生する前の12ヶ月間にクレームの原因となったサービスに対してお客様が支払った金額と100ドル（100円）のいずれか大きい方を超えないものとします。本条に定める制限は、適用される法律で認められる最大限の範囲にのみ適用されます。

[Translated with DeepL](#)

8. 紛争解決

お客様は、以下の強制的仲裁および集団訴訟放棄の規定に同意するものとします：

(a) 強制的な仲裁。ただし、お客様は、本仲裁条件または関連する変更に同意してから 30 日以内に本フォームに記入することにより、本仲裁条件および本仲裁条件の今後の変更からオプトアウトする権利を有します。

(b) 非公式な紛争解決。私たちは、正式な法的措置の前に、あなたの懸念を理解し、対処するよう努めたいと考えています。OpenAI に対して申し立てを行う前に、お客様は、お客様の氏名、紛争の内容、およびお客様が求める救済を dispute-resolution@openai.com に通知することによって、紛争の非公式な解決を試みることに同意するものとします。60日以内に紛争を解決できない場合、あなたは正式な手続きを取ることができます。時効は、60日間の解決プロセス中に停止します。お客様がEUに居住している場合、欧州委員会は、オンライン紛争解決プラットフォームを提供しており、お客様は、<https://ec.europa.eu/consumers/odr>、これにアクセスすることができます。

(c) 仲裁の場。いずれの当事者も、代替紛争解決プロバイダーであるADRサービスを通じて拘束力のある仲裁を開始することができます。当事者は、仲裁手数料を均等に負担します。仲裁人が、あなたが仲裁費用を支払う余裕がなく、放棄を得ることができないと判断した場合、OpenAIがあなたに代わって仲裁費用を支払います。OpenAIは、仲裁人がお客様の請求が軽薄であると判断しない限り、仲裁において弁護士費用および経費を求めません。

(d) 仲裁の手続き。仲裁は、カリフォルニア州サンフランシスコ、または相互に合意した別の場所において、電話、書面による提出に基づく、ビデオ会議、または直接会って行われるものとします。仲裁は、ADRサービスがその時点で有効な規則に基づき、単独の仲裁人によって実施されます。すべての問題は仲裁人が決定するものとしますが、ただし、カリフォルニア州の裁判所は、(i)以下の大量提出手続きを含む本第8条の範囲、執行可能性および仲裁可能性、ならびに(ii)お客様が本条における仲裁前の要件を遵守しているかどうかを判断する権限を有します。和解案の金額は、仲裁人が最終裁定を決定する場合、その決定後まで、いずれの当事者も仲裁人に開示しないものとします。

(e). 例外事項。(i) 少額裁判所で提起された個人の請求、および(ii) 本サービスの無許可の使用もしくは悪用または知的財産の侵害を阻止するための差止命令による救済またはその他の衡平法上の救済。

(f) 集団訴訟の禁止。紛争は個人単位でのみ提起されなければならない、集団訴訟、連結訴訟、または代表訴訟と称する訴訟において原告または集団メンバーとして提起することはできません。集団仲裁、集団訴訟、私選弁護士訴訟、および他の仲裁との併合は認められません。何らかの理由で紛争が仲裁ではなく法廷で進行した場合、各当事者は、故意に、かつ取消不能形で、あらゆる訴訟、手続き、または反訴において陪審員による裁判を受ける権利を放棄するものとします。これは、いずれの当事者も、請求のクラス全体の和解に参加することを妨げるものではありません。

(g) 集団提訴。いつでも、30以上の同様の仲裁要求が、同一または協調する弁護士または団体によってOpenAIまたは関連当事者に対して主張される場合（以下「大量申立」）、ADRサービスは、大量申立のそれぞれにランダムに連番を割り当てます。1〜10番のクレームは「初期テストケース」となり、最初に仲裁に進むこととなります。仲裁人は、クレームが事前に解決されるか、または当事者が期限を延長することに合意しない限り、最初のヒアリング前会議から120日以内にイニシャルテスト・ケースの最終裁定を下します。その後、当事者は90日間（以下「調停期間」）、初回テストケースの裁定に基づき、残りのケースを調停で解決することができます。この期間中に未解決の請求を解決できない場合、当事者は、調停期間後60日以内に相手方に書面で通知することにより、仲裁手続から脱退して法廷で手続を行うことを選択できます。そうでない場合は、残りの案件は、割り当てられた順序で仲裁されます。いかなる時効も、最初のテストケースが選択された時点から、お客様のケースが上記のように選択されるまで、停止されるものとします。

(h) 分離可能性。ただし、部分的な違法性または執行不能性の認定により、大量提出または集団もしくは代表者による仲裁が認められる場合は、本第8条の全体が執行不能となるものとします。本条項のいかなる規定も、仲裁人から当該請求の実体に関する裁定を受けるまでの間、公的な差止救済を求める権利またはその他の放棄できない権利を放棄またはその他の方法で制限するとみなされることはないものとします。

9. 一般条項

(a) 両当事者の関係 本規約は、お客様とOpenAIまたはOpenAIの関連会社との間に、パートナーシップ、ジョイントベンチャー、代理店関係を構築するものではありません。OpenAI とお客様は独立した契約者であり、いずれの当事者も、相手方の書面による事前の同意なしに、相手方を拘束したり、相手方のために義務を負ったりする権限を有さないものとします。

(b) ブランドの使用。お客様は、当社の事前の書面による同意なしに、OpenAI またはその関連会社の名前、ロゴ、または商標を使用することはできません。

(c) 米国連邦政府機関事業者。本サービスは、私費のみで開発されたものであり、適用される米国連邦調達規則およびその機関の補足の意味において、商用コンピュータソフトウェアおよび関連文書です。

(d) 著作権に関する苦情。お客様の知的財産権が侵害されたと思われる場合は、以下の住所に通知を送るか、このフォームに必要な事項を記入してください。当社は、侵害の疑いがあるコンテンツを削除または無効にし、侵害を繰り返す人のアカウントを停止することがあります。

OpenAI, L.L.C.

3180 18th St

サンフランシスコ, CA 94110

担当：ゼネラルカウンセラー/著作権代理人

著作権侵害に関する書面による請求には、以下の情報を含める必要があります：

著作権の権利者の代理として行動することを認められた人物の物理的または電子的な署名；

著作権が侵害されたと主張する著作物の説明；

侵害されていると主張する著作物がサイト上のどこにあるのかの説明；

あなたの住所、電話番号、および電子メールアドレス；

著作権所有者、その代理人、または法律により、係争中の使用が許可されていないことを誠実に信じるというお客様の声明。

偽証罪に問われることを前提に作成された、通知における上記の情報が正確であり、お客様が著作権者であること、または著作権者の代理として行動する権限を有することを示すお客様の声明。

(e) 譲渡および委任。お客様は、支配権の変更に関連する場合を含め、本規約に基づきいかなる権利または義務も譲渡または委任することはできません。譲渡および委任と称するものは、無効とします。当社は、当社の資産の全部または実質的に全部の合併、買収または売却に関連して、または関連会社に対して、または企業再編の一環として、本規約を譲渡することができます。

[Translated with DeepL](#)

- (f) 改定。当社は、本規約の改訂版をウェブサイトに掲載することにより、本規約を随時変更することができ、また、本規約に基づくお客様の権利または義務に重大な悪影響を及ぼす更新がなされた場合、おお客様に通知する電子メールまたはおお客様に通知する電子メールを送付するものとします。おお客様は、当社がお客様に通知する後、30日以内に、有効となるものとします。その他の変更は、直ちに有効となり、変更後、おお客様が本サービスを継続して使用することは、おお客様が当該変更にご同意したことを意味します。

(g) 通知。すべての通知は、書面により行われます。当社は、お客様が提供した登録情報またはお客様のサービスの使用に関連する電子メールアドレスを使用してお客様に通知することができます。送達には、電子メールで送達された場合は受領日、郵便で送達された場合は宅配便で送られた日に行われたとみなされます。OpenAIは、この住所での送達を受け付けます： OpenAI, L.L.C., 3180 18th Street, San Francisco, CA 94110, Attn: contract-notices@openai.com.

(h) 権利放棄と分離性。お客様が本規約を遵守せず、OpenAIがすぐに行動を起こさない場合でも、これはOpenAIが当社の権利を放棄していることを意味するものではありません。第8条に定める場合を除き、本規約の一部が管轄裁判所によって無効または執行不能と判断された場合、その条項は許容される最大限の範囲内で執行され、他の条項の執行可能性に影響を与えることはありません。

(i) 輸出管理。本サービスは、(a)米国の禁輸国（以下、総称して「禁輸国」）において、または(b)米国財務省の禁輸国リストに記載されている者に対して、使用またはその利益のために輸出または再輸出してはなりません。米国財務省の特別指定国民リスト、米当業者リスト（現在または将来の外資企業）、米商務省の拒否者リスト（現在または将来の団体）、またはその他の制限付き当業者リスト（以下、総称して「制限付き当業者リスト」といいます。） おお客様は、おお客様が禁輸国に所在せず、かかる制限された当業者リストに掲載されていないことをおお客様が保証します。おお客様は、エントリを直接知るための要件または義務を含め、禁輸国または制限付き当業者リストに関連するすべての適用法を遵守しなければなりません。

(j) 衡平法上の救済措置。お客様は、お客様が本規約に違反した場合、OpenAIおよびその関連会社に回復不能な損害を与える可能性があることを認め、OpenAIは、他の法的救済に加えて、お客様に対して差し止めによる救済を求める権利を有するものとします。

(k) 完全合意。本規約および本規約に組み込まれたポリシーは、本サービスの使用に関するお客様とOpenAIとの間の完全な合意を含み、サービス固有の使用条件または適用される企業契約を除き、その主題に関するおお客様とOpenAIとの間の過去または同時のあらゆる合意、通信または了解に取って代わります。

- (l) 裁判管轄、裁判地、および法の選択。本規約は、カリフォルニア州の抵触法の規則または原則を除き、カリフォルニア州の法律に準拠するものとします。紛争解決セクションで規定されている場合を除き、本規約に起因または関連するすべての請求は、米国カリフォルニア州サンフランシスコ郡の連邦裁判所または州裁判所に専属的に提起されるものとします。

- [Translated with DeepL](#)

その他のポリシー有り

Usage Policy

<https://ascii.jp/elem/000/004/130/4130807/>

etc.

別の回答から発展した対話例

プロンプト：「ChatGPT を社内で活用する際のリスクを考え、リスクを最小化するためのガイドラインやルールを考案して定めておく必要があると思います。機密情報や個人情報を送信しないのはもちろん、生成結果のファクトチェックや適切さの判断、念のために検索して第三者の著作権に抵触しないかなどがあるでしょう。日本の著作権法や、OpenAI社の利用規約などを含め、ガイドラインに盛り込むべきことを、10箇条ほど箇条書きにしてください。」



• Bing バランスモード

- 送信済みメッセージ. ChatGPT を社内で活用する際のリスクを考え、リスクを最小化するためのガイドラインやルールを考案して定めておく必要があると思います。機密情報や個人情報を送信しないのはもちろん、生成結果のファクトチェックや適切さの判断、念のために検索して第三者の著作権に抵触しないかなどがあるでしょう。日本の著作権法や、OpenAI社の利用規約などを含め、ガイドラインに盛り込むべきことを、10箇条ほど箇条書きにしてください。
- ChatGPTを社内で活用する際には、リスクを最小化するためのガイドラインやルールを定めることが重要です。以下に、10箇条ほど箇条書きにして、盛り込むべきことを挙げてみました。
 1. 機密情報や個人情報を送信しないこと。
 2. 生成結果のファクトチェックや適切さの判断を行うこと。
 3. 念のために検索して第三者の著作権に抵触しないか確認すること。
 4. 日本の著作権法や、OpenAI社の利用規約などを確認し、遵守すること。
 5. ユーザーからの要求に対して、適切な返答を行うこと。
 6. ユーザーからの要求に対して、適切な情報提供を行うこと。
 7. ユーザーからの要求に対して、適切なアドバイスを行うこと。
 8. ユーザーからの要求に対して、適切なサポートを行うこと。
 9. ユーザーからの要求に対して、適切なフィードバックを行うこと。
 10. ユーザーからの要求に対して、適切なアクションを行うこと。
- 以上が、ChatGPTを社内で活用する際に盛り込むべきガイドラインやルールです。ご参考までにお役立てください。



- 上記で「ユーザー」とはChatGPTを利用する社員のことですか？ その「ユーザー」に適切なアドバイスやサポート、フィードバックを行うのは、情報システム部のAI利用サポート担当者ということですか？
- はい、その通りです。ここでの「ユーザー」とは、ChatGPTを利用する社員のことを指します。また、適切なアドバイスやサポート、フィードバックを行うのは、情報システム部のAI利用サポート担当者が担当することになります。
- Answer from this page instead
- What is the role of the AI support staff?
- What is the purpose of ChatGPT?
- How can I contact the AI support staff?



- 情報システム部のAI利用サポート担当者が不足する合に、その一部をChatGPT自身で代替させることはできますか？
- ‘ChatGPTが情報システム部のAI利用サポート担当者を代替できるかどうか’を検索しています
- 回答を生成しています...
- ChatGPTは、情報システム部のAI利用サポート担当者を代替することはできません。ChatGPTは、人工知能チャットボットであり、自然言語処理技術を使用して、人間との対話をシミュレートすることができますが、AI利用サポート担当者のような役割を完全に代替することはできません。ただし、ChatGPTは、一部のタスクや問題に対して有用な情報を提供することができます。
- 詳細情報:[1. www3.nhk.or.jp](http://www3.nhk.or.jp)[2. ja.wikipedia.org](http://ja.wikipedia.org)[3. japan.cnet.com](http://japan.cnet.com)+2 その他
- Answer from this page insteadChatGPTとは何ですか？ChatGPTはどのように機能しますか？ChatGPTはどのようなタスクに対応できます



How can employees use ChatGPT responsibly?

- The current web page context is empty. Here are some tips for using ChatGPT responsibly within a company:
 1. Understand the capabilities and limitations of ChatGPT.
 2. Use ChatGPT for appropriate tasks and in accordance with company policies and guidelines.
 3. Do not use ChatGPT to generate or transmit confidential or personal information.
 4. Fact-check the generated results and evaluate their appropriateness before using them.
 5. Be mindful of potential biases in the generated content and take steps to mitigate them.
 6. Report any issues or concerns to the appropriate person or department within the company.
- By following these tips, employees can use ChatGPT responsibly and effectively within a company setting. Is there anything else you would like to know?
- By following these tips, employees can use ChatGPT responsibly and effectively within a company setting. Is there anything else you would like to know??

社員が責任を持ってChatGPTを利用するには？

現在のWebページのコンテキストは空です。ChatGPTを社内で責任を持って使用するためのヒントを紹介します：

- 1.ChatGPTの機能と制限を理解する。
- 2.ChatGPTの機能と限界を理解し、会社のポリシーとガイドラインに従って、適切なタスクにChatGPTを使用します。
3. ChatGPTを機密情報や個人情報の生成や送信に使用しない。
4. 生成された結果をファクトチェックし、その妥当性を評価した上で使用すること。
5. 生成されたコンテンツに潜在する偏見に留意し、それを軽減するための手段を講じること。
- 6.問題や懸念がある場合は、社内の適切な担当者または部署に報告する。

これらのヒントに従うことで、従業員は会社で責任を持って効果的にChatGPTを使用することができます。他に何かお知りになりたいことはありますか？

これらのヒントに従うことで、従業員は会社で責任を持って効果的にChatGPTを使用することができます。他に何か知りたいことはありますか？

[Translated with DeepL](#)

社員が責任を持ってChatGPTを利用するためには？ 2023.5.27



ChatGPTを社内で責任を持って使用するためのヒントを紹介します：

1. ChatGPTの機能と制限を理解する。
2. ChatGPTの機能と限界を理解し、会社のポリシーとガイドラインに従って、適切なタスクにChatGPTを使用します。
3. ChatGPTを機密情報や個人情報の生成や送信に使用しない。
4. 生成された結果をファクトチェックし、その適切さを評価した上で使用すること。
5. 生成されたコンテンツに潜在する偏見に留意し、それを軽減するための手段を講じること。
6. 問題や懸念がある場合は、社内の適切な担当者または部署に報告する。

これらのヒントに従うことで、従業員は会社で責任を持って効果的にChatGPTを使用することができます。

ChatGPTなどを正しく使うためには、その分野の知識をある程度もち、なぜそのようになってるかを判断する論理的思考力を強く働かせるのに加えて、"暗黙の前提"を排除した質問スタイルを取ることが大切。

社員が責任を持ってChatGPTを利用するためには？ 2023.6.13



- ChatGPTを責任を持って利用するためには、以下のガイドラインに従うことが重要です。
- 1. プライバシーとセキュリティの尊重: ChatGPTを使用する際には、個人情報や機密情報などのプライバシーに関わるデータを開示しないように注意しましょう。また、セキュリティ上のリスクを最小限に抑えるために、公共のネットワークや不正確なセキュリティ設定が行われたデバイスからのアクセスを避けてください。
- 2. 偏見や差別の回避: ChatGPTを利用する際には、人種、性別、宗教、性的指向、障害、国籍などに基づく偏見や差別的な表現を避けましょう。個別のユーザーに対する攻撃的な言葉遣いや侮辱も避けるようにしましょう。
- 3. 不正確な情報の拡散を避ける: ChatGPTは大量のデータから学習したモデルですが、間違った情報や誤解を招く可能性もあります。特に医療や法律などの専門知識が必要な領域では、正確な情報を提供するためには専門家の助言を求めることが重要です。
- 4. ユーザーの心理的・感情的な健康を考慮する: ChatGPTは、人間のように感情や心理的な状態を理解することはできません。したがって、敏感なトピックや個人的な問題についての助言を求める場合でも、専門家や信頼できる人々に相談することをおすすめします。
- 5. アルゴリズムの誤解を避ける: ChatGPTは、与えられたコンテキストに基づいて応答を生成する自律的なモデルですが、それが常に正確で適切な回答を提供するわけではありません。アルゴリズムの限界やバイアスを認識し、その制約を理解することが重要です。
- 6. 適切な利用範囲で使用する: ChatGPTは、一般的な情報提供や質問応答の目的で設計されています。営利目的や悪意のある活動、不法行為に利用することは避けましょう。
- これらのガイドラインに従うことで、社員は責任を持ってChatGPTを利用することができます。また、適切なトレーニングやガイドラインの提供、定期的なアップデートと監視など、組織全体での対応策を確立することも重要です。

社員が責任を持ってChatGPTを利用するためには？ 2023.6.13



- ChatGPTを責任ある方法で使用するためには、以下のガイドラインを考慮してください。
- 1. **ユーザープライバシーの尊重:** ChatGPTがユーザーから収集したデータは機密性を持ち、個人情報の保護が必要です。ユーザーから取得した情報は適切な許可なく共有したり、使用したりしないでください。
- 2. **適切な使用:** ChatGPTは、不適切または違法な行為に使用するべきではありません。これには、誹謗中傷、ハラスメント、違法な情報の共有などが含まれます。
- 3. **信頼性の理解:** ChatGPTは人間が作成した情報を基に学習を行いますが、常に100%正確な情報を提供するわけではありません。そのため、重要な決定を下す際には、ChatGPTの情報を単独で使用するのではなく、他の信頼性の高い情報源をもとに行うべきです。
- 4. **AIの限界の理解:** ChatGPTは、質問の文脈を理解し、関連する情報を生成することができますが、個々のユーザーや特定の状況についての個別の知識を持っているわけではありません。また、モデルの知識は訓練データまでのものであり、それ以降の情報については知らないという限界があります。
- 5. **フィードバックと改善:** ChatGPTは完全ではなく、ミスをすることがあります。そのため、ユーザーからのフィードバックは非常に重要で、それによりモデルの改善と発展が進められます。
- これらのガイドラインは、ChatGPTを安全に、効果的に、そして責任ある方法で使用するためのものです。

ChatGPTの機能と限界は何ですか？



- ChatGPTは、OpenAIが開発した自然言語処理技術で、その優れた機能から大きな注目を集めています1. 既存のデータから学習したパターンや構造に基づいて新しいデータを作成することができる生成型AIモデルです2. そのため、テキスト、画像、音楽など、さまざまな領域にまたがるコンテンツを生成することができます 2.

しかし、ChatGPTにはいくつかの限界もあります。その一つは、既存のデータやアルゴリズムに依存していることです3. ChatGPTはユーザーのインタラクションから学習することができますが、最終的には学習させたデータによって制限されます3. つまり、ChatGPTは、トレーニングデータ以外の馴染みのないトピックやクエリに苦戦し、不正確な回答や役に立たない回答につながる可能性があります3. さらに、ChatGPTは、偏ったデータ、疑わしいコンテンツ、事実上不正確な情報に満ちている可能性のある公開データのみでトレーニングされているため、偏った、または不正確なコンテンツを生成する可能性があります 45.

詳細情報

1. researchgate.net
2. sciencedirect.com
3. minitool.com (ミニツール・ドット・コム)
4. computerweekly.com
5. chat-gpt.tools

Questions:

- ChatGPTのアプリケーションにはどのようなものがありますか？
- 生成されたコンテンツの潜在的なバイアスを軽減するにはどうすればよいですか？
- 生成されたコンテンツに含まれる潜在的なバイアスにはどのようなものがありますか？

ChatGPT API オプトアウトの方法

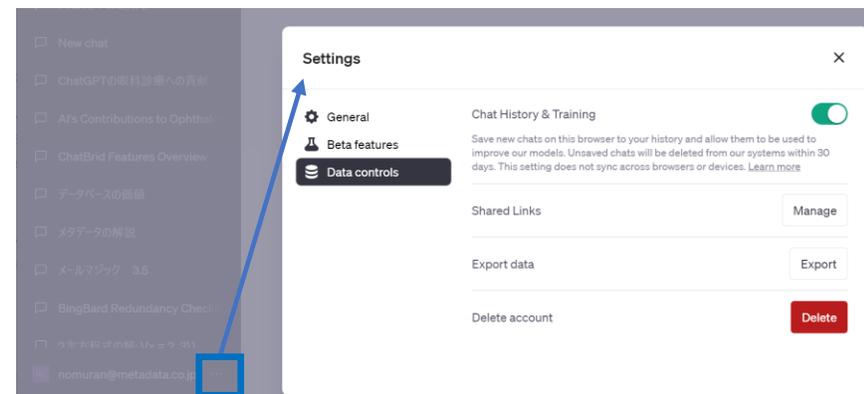
~ChatGPT APIに入力情報を学習データとして使用させない方法「オプトアウト」。

ChatGPT APIに対して、入力情報を学習データとして使用しないように指示することができます。これにより、入力情報が他のユーザーに影響を与えることを防ぐことができます。

オプトアウトの方法を詳しく教えてください。

‘ChatGPT API オプトアウトの方法’を検索しています

- オプトアウトの方法は、以下のようになります12。
 - ChatGPTのオプトアウトリクエストフォーム1を開く
 - メールIDを使用してにログインする
 - フォームに詳細を入力する
 - 送信をクリックする
- これで、ChatGPT APIに入力情報を学習データとして使用させないことができます。ただし、オプトアウトすると、性能に制限される場合があることに注意してください1。
- 詳細情報: 1. bing.com 2. learn.microsoft.com 3. qiita.com 4. gamingdeputy.com



- オプトアウトした場合の性能の制限はどのようなものですか？
‘ChatGPT API オプトアウト 性能’を検索しています
- オプトアウトした場合の性能の制限は、以下のようになります1。
 - ChatGPTやDALL-E2を使用する際、提供されたデータを使用することがあるが、オプトアウトするとそのデータが使用されなくなる。そのため、回答の精度が低下する可能性がある。
- 以上のようにして、オプトアウトした場合の性能の制限を教えました。オプトアウトするかどうかは、自分の用途やニーズに応じて判断する必要があります。
- 詳細情報: 1. blog.serverworks.co.jp 2. qiita.com 3. qiita.com

あるいは、[この文書によれば](#)、オプトアウトする代わりに、左下のアカウント名の右をクリックして設定Settingsを開きData controlsでChat History & Trainingをoffにしても学習されない。

でも、左カラムに履歴が残ると大変便利なので、オプトアウトしましょう！

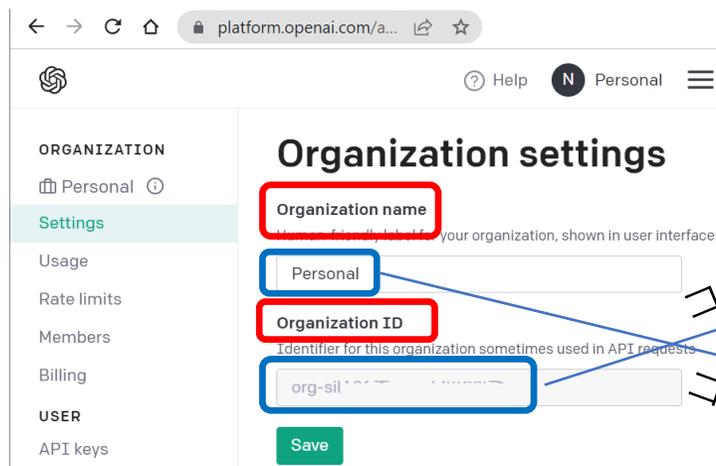
今すぐオプトアウトしましょう！

- Data usage for consumer services FAQ

- <https://help.openai.com/en/articles/7039943-data-usage-for-consumer-services-faq>
- この最初の段落の5行目あたりにある下線付き [form](#) をクリックしてください。そこで開く下記のページを1分くらいで埋めたら完了です。Organization IDなどは、続く括弧内の [Account Org Settings](#) をクリックしたら開きます。グレーアウトされていますが、コピペできますので、コピーバッファに入れて、下記フォームにペーストしてください。

- "User Content Opt Out Request"

- <https://docs.google.com/forms/d/1t2y-arKhcjlKc1I5ohI9Gb16t6Sq-iaYbVFEbLFFjal/viewform>



User Content Opt Out Request

One of the most useful and promising features of AI models is that they can improve over time. We continuously improve the models that power our services, such as ChatGPT and DALL-E, via scientific and engineering breakthroughs as well as exposure to real world problems and data.

As part of this continuous improvement, when you use ChatGPT or DALL-E, we may use the data you provide us to improve our models. Not only does this help our models become more accurate and better at solving your specific problem, it also helps improve their general capabilities and safety.

We know that data privacy and security are critical for our customers. We take great care to use appropriate technical and process controls to secure your data. We remove any personally identifiable information from data we intend to use to improve model performance.

We understand that in some cases you may not want your data used to improve model performance. You can opt out of having your data used to improve our models by filling out this form. Please note that in some cases this will limit the ability of our models to better address your specific use case.

For details on our data policy, please see our [Privacy Policy](#) and [Terms of Use](#) documents.

Please ensure the email you provide is associated with [your account](#), and that the **Organization ID is of the format "org-eXam3pleOr9giD" otherwise we will not be able to process your request.*

Account Settings > Account Settings > アカウントを切り替える

* 必須の質問です

メールアドレス *

メールアドレス

Organization ID (found on [Account Org Settings](#)) *

回答を入力

Organization name (found on [Account Org Settings](#))

回答を入力

回答のコピーが指定したアドレスにメールで送信されます。

Data usage for consumer services FAQ の他の重要ポイント

コンシューマ向けサービスのデータ利用に関するFAQ

1. (前ページ) オプトアウト申請後、アプリでの入力データは学習を停止。
2. OpenAI社のAPI製品については、顧客(ユーザ)が明示的に入力データをOpenAI社と共有しようとして決定し(てアクションし)ない限り入力データを学習には使いません。
3. オプトアウトしていない場合、例えば[このように](#)入力データはOpenAIサービスの改良に使われます。
4. [こちらの手順](#)で、アカウントを削除できる。削除依頼を送信してから30日以内にアカウントを削除します。セキュリティ上の理由から、一度アカウントを削除すると、同じメールアドレスでアカウントを再登録不可。
5. すべてのチャット履歴を永久に削除したい場合、チャット履歴から特定のチャットの会話を1回ずつ消去する(しかない)。30日後、クリアされた会話は永久に削除されるか、または、当社のサービス向上のためにお客様のコンテンツを使用することをお客様がオプトアウトしていない場合は、非特定化(誰の入力が分からない形に)されます。
6. 「私の作品は第三者と共有されるのでしょうか？」
 1. 「当社は、当社のサービス提供を支援する信頼できる[サービスプロバイダーの一部](#)とコンテンツを共有しています。私たちは、この目的を達成するために必要な最小限のコンテンツを共有し、私たちのサービスプロバイダーは、厳格な機密保持とセキュリティの義務に従うものとします。」→恐らく他社クラウドに置いているとの意味。
 2. 「当社は、マーケティングや広告の目的でユーザーコンテンツを使用または共有することはありません。」
7. 「人間が私のコンテンツを閲覧するのですか？」
 1. 限られた人数のOpenAIの正規職員、および機密保持とセキュリティの義務を負う専門的な第三者請負業者は、以下の理由で必要な場合にのみ、ユーザーコンテンツを閲覧し、アクセスすることができます：(1) 不正使用やセキュリティ事故の調査、(2) アカウントに関する質問で当社に連絡した場合のサポート提供、(3) 法的義務の遵守、(4) ユーザー投稿データを使用してモデルを微調整する場合(あなたがオプトアウトしていない限り)、当社はPIIフィルター技術を使用して個人データの使用量を削減することもあります。コンテンツへのアクセスは、技術的なアクセスコントロールの対象となり、知る必要のある権限のある担当者だけに制限されます。さらに、当社は、ユーザーコンテンツへのすべてのアクセスを監視し、ログに記録し、権限を与えられた担当者は、ユーザーコンテンツにアクセスする前に、セキュリティおよびプライバシーのトレーニングを受けなければなりません。

[Translated with DeepL](#)

付録

深層学習

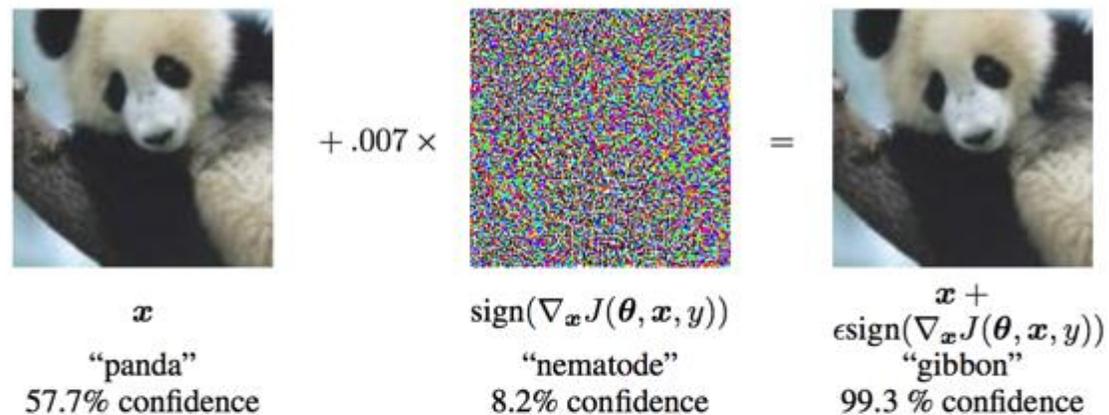
ディープラーニング特有のセキュリティ問題

～主に画像を例に

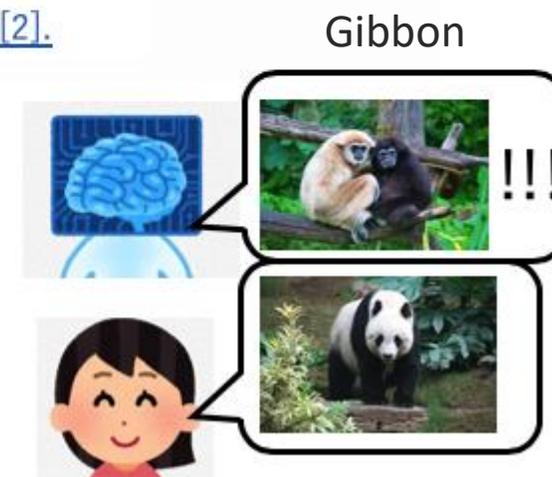
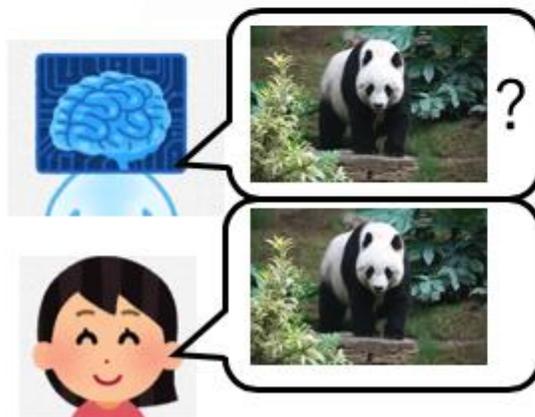
- Adversarial Example: 判定の際のテストデータの改ざんによる誤判定。
- Training Set Poisoning: トレーニングデータに、悪意ある誤判定させる目的のデータを混ぜること。
- [Pro] Trojan Horse: AI応用システムが活用するニューラルネットワークの構造自体に、将来、判定全体がゆがむような仕組みを内包させること。

1 「敵対的サンプル画像 (Adversarial Examples)」

- 正しいラベルに分類したときにロスが大きくなるような摂動を加え加工し誤認へ
- Adversarial Attackやそれを防御する方法は、現在盛んな研究分野の1つ
- “悪意ある攻撃として、まず認識や判定の際に、人間が見たら元画像のように見える画像にAIだけをだます特殊な加工を施す手法があります。これは「敵対的サンプル画像(Adversarial Examples)」と呼ばれ、100%に近い確信度で全然違う判定結果に導かれます。
- 深層学習は非線形で、画像認識では多対一のマッピングです。このため原理的に、ある分類カテゴリーに似ても似つかない元画像を落とし込むことができず、ある程度、総当たりの試行錯誤の上、ある種の画像加工のパターンを見つけられなかったら、時には人命に関わる重大な誤判定を引き起こされてしまいます。今後1億シーンを学習させるとも言われる自動運転車向けに、障害物をそうでないと思わせたり、交通標識を正反対の意味のものに意図的に誤認させるステッカーを貼られたりしたら、苦労して実用可能な精度にまで進化させた自動運転車の取り組みが再び振り出しに戻ることになりかねません。”
 - 最強のAI活用術, 野村2017)



[Goodfellow et al. \(2014\)\[2\].](#)



Adversarial Examples Improve Image Recognition 2019 末 日本語解説記事

- GAN (敵対的生成ネットワーク) と同様の発想で、攻撃をかわせるように鍛えることも出来るという発想で解決へ
 - “EfficientNetは2019年5月に提案された、ネットワーク自動探索 (Neural Architecture Search) で最適化されたネットワーク”

ImageNet-A[4]

Black: Actual Class
Red: ResNet50 prediction class and confidence.



ImageNet-C[5]

corrupted ImageNet



Stylized-ImageNet[6]

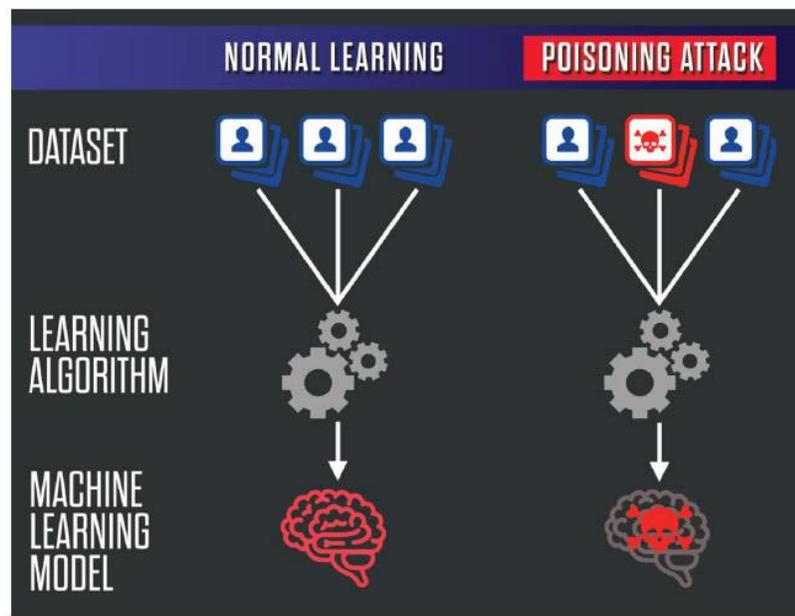
created by applying AdaIN style transfer to ImageNet images.



2 「トレーニングセットへの毒仕込み (Training Set Poisoning)」

- “もう一つは、AI開発のインサイダーによる犯罪、もしくは開発者の内部ネットワークに侵入して、大量のトレーニングデータに意図的にノイズを紛れ込ませる手法です。「トレーニングセットへの毒仕込み (Training Set Poisoning)」などと呼ばれます。ライバル企業のAI開発を失敗させる、ないしは遅らせる、あるいは業界へのAI導入自体を阻んで既得権を温存したいとする勢力が、今後このような犯罪に手を染める可能性が無きにしもあらずでしょうか。”

- 最強のAI活用術, 野村2017



<https://informationmatters.net/data-poisoning-ai/>

3 「トロイの木馬 (Trojan Horse)」

<https://arxiv.org/pdf/1902.06531.pdf>

- データポイズニングの発展形。
 - 比喩的に言えばWebへのSQLインジェクション攻撃に似ている
 - トロイの木馬は、DNNモデルの解釈の難しさを利用して、DNNモデルに作られた有効なバックドアを利用
 - トロイの木馬攻撃は、DNNモデルに有効なバックドアを設け、学習したモデルの解釈の難しさを利用して、攻撃者が選んだトロイの木馬のトリガーを選択
 - →どれがトリガーかは通常、被害者には全くわからない！
- ”[STRIP: A Defence Against Trojan Attacks on Deep Neural Networks](#)“
 - 意図的に入力を乱し、乱された入力に対する予測クラスのランダム性を観察
 - 予測クラスのエントロピーが低いと、良性モデルの入力依存性に反し、悪意のある入力の存在が示唆される

ChatGPTで懸念されるセキュリティ攻撃

- プロンプト・インジェクションによる脱獄=Jail Break
 - 本来、膨大なコストをかけて排除してきた有害コンテンツを吐き出させるようなプロンプト生成のテクニック
 - 爆弾、麻薬の製造方法、公序良俗に反するHowToなどを吐き出させるテクニックは存在し、比較的容易にヒットさせることができる。
 - プロンプトエンジニアリングのテクニックに長けていれば、生成系AIに（途中まで）白状させることもできる。
- →分類、実例、対策は[こちら](#)：
 - 【徹底解説】今知るべきChatGPTの「脱獄」手法による攻撃とは？
「Macnica Security Forum 2023」 凌翔太氏講演レポート

プロパガンダの見破り方： [ChatGPT時代のAI講座](#) →

- <https://www.youtube.com/watch?v=-gUMiuoFCUE>

ChatGPT4にプロパガンダの見破り方を訊!

ある記事が、情報発信者や第三者のプロパガンダであると見破るコツを7つ教えてください。

プロパガンダを見破るために、以下の7つのコツを試してみてください。

1. 情報源を確認：情報発信者が信頼性のある組織や個人かどうか調べてください。不明瞭な情報源や偏激なウェブサイトには注意が必要です。
2. 情報の正確性を検証：提示されたデータや事実について、独立した信頼性の高い情報源で確認してください。一貫性のない情報や誇張された主張には注意してください。
3. 言語とトーンに注意：感情的な言葉や極端な表現が使われている場合、プロパガンダである可能性が高いです。客観的なトーンで書かれた記事には、通常、信頼性があります。
4. 論理的整合性を評価：記事が一貫した論理構造を持っているかどうかを確認してください。矛盾や論理の飛躍がある場合、それはプロパガンダの可能性があり得ます。
5. 偏りを見出す：記事が特定のグループや意見に偏っていないかを評価してください。バランスの取れた報道は、異なる視点や意見を取り上げることが一般的です。
6. 感情的な操作に注意：プロパガンダは、恐怖や怒りなどの感情を煽ることがよくあります。記事が感情に訴える言葉やイメージを使用している場合、その背後にある目的を検討してください。
7. 複数の情報源を比較：同じトピックに関する複数の情報源を参照して、異なる視点と比較検討してください。これにより、プロパガンダの要素が含まれているかどうかを判断しやすくなります。

これらのコツを使用して、プロパガンダの痕跡を見つけ出すことができます。ただし、情報を検証し、偏りや感情的な操作に対する能力を養うことも重要です。常に懐疑的な目で情報を見て、客観的な判断を下すよう努めてください。

ありがとう。この7つのコツを、小学3年生が理解できるように書き直してください。

小学3年生向けに、プロパガンダを見破る7つのコツを簡単に説明します。

1. だれが言っているのか調べる：話をしている人や組織が信頼できるかどうかを見てください。
2. 本当かどうか確かめる：言われていることが本当かどうか、ほかの信頼できる場所で調べてみましょう。
3. 言葉や声の調子に注意：怒ったり、悲しんだりするような言葉がたくさん使われていたら、気をつけましょう。
4. 話がつじつまが合っているか見る：話がうまくつながっていて、おかしいところがないか確かめましょう。
5. どちらかの意見ばかり言っていないか見る：いろんな意見や考え方が言われているかどうか、チェックしましょう。
6. 気持ちをあおられていないか注意：怖いことや怒りを感じさせる言葉が使われていたら、その理由を考えてみましょう。
7. いろんなところで調べる：同じ話題について、たくさんの情報を集めて比べてみましょう。

これらのコツを使って、うそや変な話を見つけることができます。いつも注意して、正しいことを見つける力を身につけましょう。

野村 直之氏: ChatGPT時代のAI講座

日本電子出版協会
チャンネル登録者数 1060人

4481 回視聴 1か月前に配信済み
00:00 お話の流れ (目次)
04:19 1. ChatGPT # 高度な検索ツール