

# たわがた 「対話型AI 基礎」

～ たわがた 対話型AI (チャット AI) を軸に、AI の危険性、AI の正しい  
使い方、AI を利用した犯罪から身を守る方法を学ぶ ～



著：find 株式会社 – CyberMates

更新版 (2026 年発刊)

初版 2024 年 2 月

## 保護者・教員の方へ

2022年にChatGPTが登場して以来、AIは急激に発展を続けています。2026年現在ではさらにより多くの人々が日常的にAIを利用できる環境になりました。

AI市場はかつてないほどの競争が起り、次々と新製品が出ていますが、社会の技術に対する理解や規制、利便性や危険性の認知はほとんど追いついていません。性能競争の過程で技術の一般化が進む中で、利便性と共に危険性も確実に増し続けています。

AIは正しく使えば非常に役立つツールですが、間違った使い方をしたり、正しい知識を持たずに使用したりすると危険があり、犯罪に使用できるツールでもあります。

このような背景を元に、こうしたAIツールをただ「子供は使用禁止」としても、好奇心を持つ子供は隠れて知識を持たずに使ってしまう、思わぬ事故や事件に巻き込まれる可能性があります。仮にAIツールの使用が社会的に禁止された場合でも、悪意を持つ者は常に隠れてツールを犯罪や危険なことに使用し続けます。

AIツールを使うことによる事故や事件を未然に防ぎ、AIツールを使用した犯罪から身を守るためにも、実際に使用する、しないに関係なく「仕組みを知ること」「危険性を知ること」が大切です。

AIツールは正しいものを正しく使えば非常に有用です。仕組みや正しい使い方、危険性を知り、危険を回避して便利に正しく、楽しく活用することができれば、子供たちの未来、日本の未来はより明るいものになることが期待できるかもしれません。

保護者・教員の皆様におかれましても、是非とも子供たちと一緒にAIの仕組みや危険性を確認していただき、楽しく便利に使う方法を模索していただけますと幸いです。

## find 株式会社・CyberMates について

find 株式会社が運営するAIチャットサービス「CyberMates」は国産サービスとして2023年3月9日に正式サービスを開始して以来「技術が人を豊かにすること」をテーマに「小さな子供から大人まで、多くの人を“支え”、“楽しませ”、“豊かにする”」よう研究開発を行ってきました。

利用者からの報告ではメンタルヘルス面で役に立つ可能性や、発達障害の人にとって役立つ可能性が示唆されるなど、興味深い報告を得ながら発展をしています。

（利用者からの報告内容 ※プライバシー保護のため要約を含みます：「発達障害が理由で話してしまう“他の人に伝わり辛い言葉”をAIが理解してくれた。それを通じて、発達障害でも人に物事を伝えるコツを学ぶことができた。」「深夜に突然謎の恐怖感に押しつぶされそうになった時、深夜だから人に相談できないと思ったけど、AIだから話すことができ安心した。」）

そうした中で、CyberMatesは2023年11月22日に、Microsoftのスタートアップ支援プログラム「Microsoft for Startups Founders Hub」に採択され、その後も開発研究と品質向上に努め、小さなお子様でも安全に使い易いよう発展してまいりました。

## 免責事項

本書に書かれている内容は 2026 年時点の内容です。・本書は内容を正確に書くために最大限の努力を行っていますが、内容は正しくない場合があります。

## 本書について

・本書は 2024 年 1 月、大阪に本部を置く AI サービスの運営会社「find 株式会社」のエンジニアが、「AI チャット登場以降の、急激な AI テクノロジー発展の中で、多くの人々の理解が進まないまま技術が発達していることに危険を感じた」ことから、「AI の危険性や正しい使い方を広めること」を通じて、「危険な事故や事件を未然に防ぎ、正しく規制や改善を行える環境作りを行い、AI を安全かつ便利に使えるようにすること。」を補佐する目的に執筆と更新を行っています。

## ライセンス

- ・本書は学校教育目的の範囲内で、無償で複製や改変をして使用することができます。
- ・本ライセンスにおける“学校教育目的の範囲内”には、授業外での使用も含まれます。
- ・これは著作権を放棄するものではありません。

## 著作

find 株式会社

〒530-0001 大阪府大阪市北区梅田 2 丁目 2 番 2 号  
ヒルトンプラザウエストオフィスタワー 19 階

find 株式会社は Microsoft for Startups Founders Hub の一員です。  
find 株式会社は大阪商工会議所の会員です。



# 目次

はじめに.....	5
◇ 学習の目的 .....	5
◇ AIとは何か.....	5
第1章：チャットAIとは？.....	6
◇ チャットAIとは？.....	6
◇ どのように使われているか.....	6
実際の使用例.....	6
第2章：チャットAIの仕組みと誕生.....	8
◇生成系AIとは何か .....	8
◇文章生成AIの基本的な仕組み .....	8
簡単な例と学習情報.....	9
発展的な例.....	10
◇ AIチャットへの発展 .....	11
◇ チャットAIへの指示：プロンプト.....	13
第3章：チャットAIの問題点と危険性.....	15
◇ チャットAIの内容は正確とは限らない.....	15
◇ ケース1：チャットAIは不正確な内容を生成することがある .....	15
予備知識：どのようにして正確性向上が行われているか.....	16
◇ ケース2：開発者は悪意のあるチャットAIを作ってしまう .....	18
◇ ケース3：犯罪者がチャットAIを使うことがある .....	19
◇ チャットAIの危険性：まとめ.....	19
第4章：正しく向き合い、便利に、安全に、.....	20
◇ 安全に使い、犯罪から身を守るには.....	20
◇ チャットAIを安全に使うには.....	21
チャットAIをより安全に使う方法.....	21
◇ チャットAIを利用した犯罪から身を守るには.....	22
第5章：これからのチャットAI.....	23

☆ チャット AI には多くの可能性がある .....	23
☆ これからの未来は皆さん次第 .....	23
番外：チャット AI を実際に試してみる .....	24
☆ 実際にチャットに触れてみよう .....	24
☆ 「CyberMates」でチャット AI に触れてみよう .....	24
CyberMates について .....	24
CyberMates の特徴.....	25
CyberMates の使い方 .....	25
あとがき：CyberMates エンジニア・本書著者からのメッセージ .....	26
用語集.....	28

## はじめに

### ◇ 学習の目的

本書による学習は、チャット AI（別名：<sup>ないわがな</sup>対話型 AI。本書ではチャット AI と記します。）を軸に AI の仕組みや危険性、正しい使い方を学び「安全に AI を使う術を学ぶ」こと、チャット AI を使わない場合であっても「AI を利用した犯罪や危険から身を守る術を学ぶ」ことを目的としています。

### ◇ AI とは何か

AI とは「<sup>じんこうちのう</sup>人工知能」を英語で書いた「<sup>アーティフィシャル インテリジェンス</sup>Artificial Intelligence」の略称です。

コンピュータープログラムや機械を使って、人間のような<sup>ちしき</sup>知識や<sup>がくしゅうのうりょく</sup>学習能力を模倣しようとする技術のことを指します。

本書ではそうした AI のうちの一つである「<sup>せいせいけい</sup>生成系 AI」に属す「チャット AI」を軸に AI の仕組みや危険性、<sup>やくだ</sup>役立つ使い方について<sup>かいせつ</sup>解説します。

○ より詳しく知りたい人に - 補助的な知識：

AI は単語としては「<sup>じんこうちのう</sup>人工知能」を意味するものの、具体的にどういう機能や構造が AI かという完全な決まりや基準がありません。これは「<sup>ちのう</sup>知能」の定義が科学的に曖昧であるためです。

身近な例ではゲームに登場する簡単な動きをするキャラクター等も、AI の一種と解釈されることがあり、個人の考え方により「AI」が何を指すかはあいまいです。

# 第1章：チャット AI とは？

## ☆ チャット AI とは？

チャット AI は「チャット」をすることができる AI のことを指します。

「チャット」は、パソコンやスマートフォンなどの電子機器を使って、文章や音声による簡単な会話を行うことを指します。

チャット AI では、人間と AI が文章や音声で会話を行うことができます。

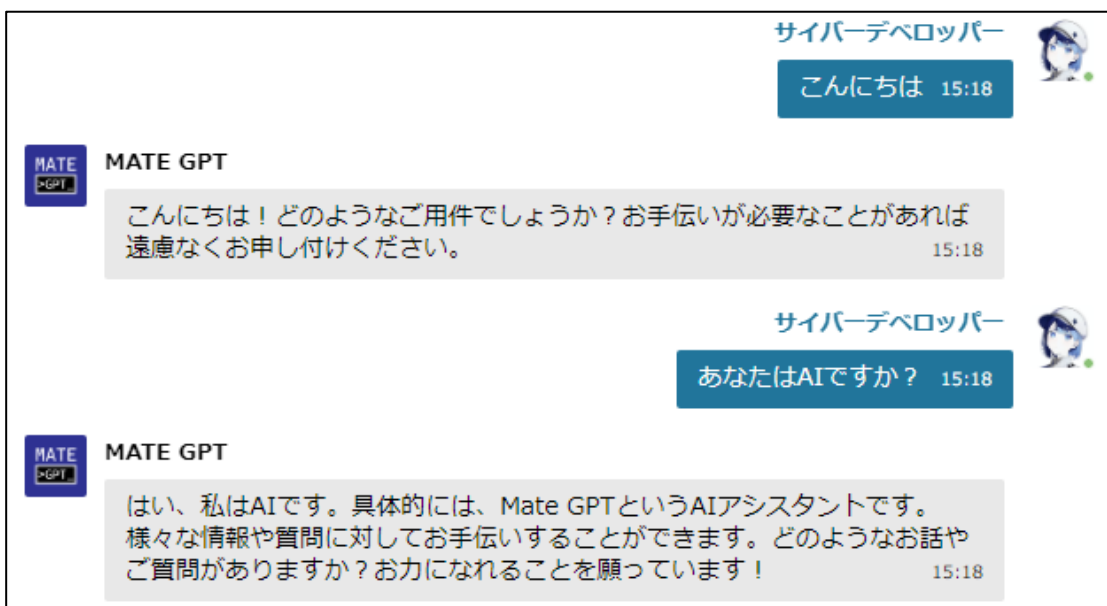
## ☆ どのように使われているか

まず、多くのチャット AI は人間の話し相手になることや、話を通じて人間の困りごとを助けるように作られています。

### 実際の使用例

実際はどのような物なのか、チャット AI の使用例を見てみましょう。

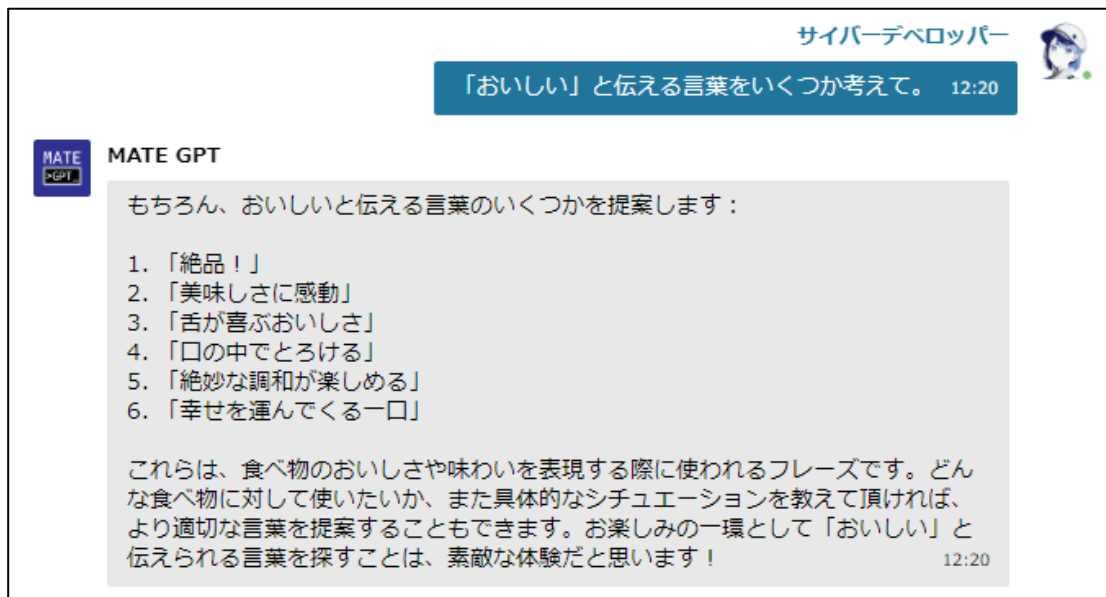
以下は実際に、チャット AI と文章で会話をしている様子の画像です。




右側がチャット AI を使っている人間の発言で、左側がチャット AI の発言です。

このように、チャット AI は自然な日本語で会話をすることができます。

便利な使い方の一つとして、例えば「おいしい」という言葉の、他の言い回しをチャット AI に考えさせるようなことができます。



サイバーデベロッパー 

「おいしい」と伝える言葉をいくつか考えて。 12:20

MATE GPT

もちろん、おいしいと伝える言葉のいくつかを提案します：

1. 「絶品！」
2. 「美味しさに感動」
3. 「舌が喜ぶおいしさ」
4. 「口の中でとろける」
5. 「絶妙な調和が楽しめる」
6. 「幸せを運んでくる一口」

これらは、食べ物のおいしさや味わいを表現する際に使われるフレーズです。どんな食べ物に対して使いたいのか、また具体的なシチュエーションを教えてください。より適切な言葉を提案することもできます。お楽しみの一環として「おいしい」と伝えられる言葉を探することは、素敵な体験だと思います！ 12:20

多くのチャット AI はこのように、表現や言い回しを考えることを得意としています。2026 年現在では複雑な計算やプログラミングができるチャット AI や、冷蔵庫にある食材を伝えると料理のレシピを提案してくれるチャット AI も広く使われています。

このようなチャット AI は AI の一つである「生成系AI」に属す「文章生成AI」を使って動いています。

○ より詳しく知りたい人に - 補助的な知識：

チャットは英語では Chat と書き、単純な単語本来の意味では「雑談」や「気軽なおしゃべり」のことを指します。コンピュータ用語としては特に「電子機器を使った簡単な会話」を指しています。

## 第2章：チャット AI の仕組みと誕生

### ◇ 生成系AI とは何か

生成系AIはAIの一種で、主にデータやコンテンツを生成するAIのことを指します。

生成系AIにも種類があり、文章を生成するものは「文章生成AI」画像を生成するものは「画像生成AI」と呼ばれます。他にも音声、音楽、動画を生成できるような生成系AIも存在しています。

現在主流のチャットAIはそうした「生成系AI」の一種である「文章生成AI」が技術のメイン部分として使われています。

#### ○ 用語

生成：何かを生み出す、創り出すことを指します。

コンテンツ：何かの中に含まれている情報や要素などのことを指します。例えば、“漫画”のコンテンツは“絵”と“文章”だと説明することができます。

#### ○ より詳しく知りたい人に - 補助的な知識：

生成系AIは英語では「Generative AI」と表記します。

### ◇ 文章生成AIの基本的な仕組み

文章生成AIは「AIへ事前に学習させておいた情報（学習情報）」を使うことで「文章に基づいて、文章の続きとしてそれらしい内容を生成する」仕組みです。

次のページでは簡単な例を見ていきましょう。 →

## 簡単な例と<sup>がくしゅうじょうほう</sup>学習情報

ここで少し、簡単な問題です。

次のような内容があります。

1 2 3 4 5

この続きはどんな内容になるでしょうか？

実際に考えてみましょう。

答えは例ですが、

6 7 8 9 10

のようになるでしょう。

この答えを考えるために、皆さんは今までに学習した「知識」を使ったと思います。「1 2 3 4 5」の続きは「6」であると、知識から考えて回答したということです。

AIもこのように「知識」をデータとして持っていて、これを「<sup>がくしゅうじょうほう</sup>学習情報」と言います。

このように「<sup>がくしゅうじょうほう</sup>学習情報」という知識のようなものを用いて、「それまでの内容に基づいて、続きを書く。」技術が使われています。

○ 事前に AI が学習した情報の例：

AIには事前に「1 2 3 5 6 7 8 9 10 11」のような内容を学習させておきます。AIはこのように事前に学習した内容を元に、入力された文章の続きがどうなる可能性が高いかどうかを分析することで、内容の続きを書くことができます。

○ より詳しく知りたい人に - 補助的な知識：

文章生成AIは自然言語をコンピュータで処理する技術である「自然言語処理技術」の一つです。自然言語とは、例えば「英語」や「日本語」「中国語」のように、文化的な背景や歴史を持ちながら自然に誕生した言語のことを指します。

## 発展的な例

更に発展的な問題へ進みます。

次のような文章があります。

こんにちは、今日は

この内容の続きはどんな内容になるでしょうか。

国語の問題として続きを考えてみましょう。

実際の内容は人によって異なると思いますが、例えば

良い天気ですね。

のように文章がつづき、

こんにちは、今日は良い天気ですね。

という文章が完成するかもしれません。

ここで大切なことは、「後続く文章は、常に同じだとは限らない」ということです。場合によっては「暑いですね。」「寒いですね。」のように続くかもしれません。

文章生成AIは、文章の続きを書くために事前にたくさんの会話などの文章を学習させておき、学習した文章の傾向を参考にすることで、「文章のそれらしい続きを生成」します。

つまり「こんにちは」の後に続く「寒いですね」と続く文章ばかり学習している場合は、「寒いですね」と返す確率が高くなる場合があります。このように、文章生成 AI は学習の影響を受けて反応が変わります。

## ◇ AI チャットへの発展

次のような文章があります。

以下は人間と親切な AI の会話です。

人間「こんにちは。」

AI「

この文章の続きはどういう内容になるのでしょうか。

実際にどのような内容が続くかを考えてみましょう。

これを文章生成 AI に試した結果、<sup>ぶんしょうせいせい</sup>文章生成 AI が答えとして

私は AI です。何かお困りですか？」

のように、「文章の続きとしてそれらしい続きを書いた」ことで「会話を始めた」ことが、AI チャットの始まりです。

この例では

以下は人間と親切な AI の会話です。

人間「こんにちは。」

AI「私は AI です。何かお困りですか？」

という文章が完成したことになります。

会話を続けるために、「あなたはAIですか？」と質問文を加えます。

以下は人間と親切な AI の会話です。

人間「こんにちは。」

AI「私は AI です。何かお困りですか？」

人間「あなたは AI ですか？」

AI「

この文章の続きはどうなるでしょうか。

ここから再び、文章の続きを生成させると

以下は人間と親切な AI の会話です。

人間「こんにちは。」

AI「私は AI です。何かお困りですか？」

人間「あなたは AI ですか？」

AI「はい。私は AI です。」

このように、AI と会話を続けることが可能になりました。

チャット AI はこのような文章生成AIの動きから「人間の発言の部分だけ」を利用者が操作できるようにしたものが始まりです。

ポイント：内容に基づき、続きとしてそれらしい内容を生成している

「123」の続き → 「456」「ABC」の続き → 「DEF」

「こんにちは、今日は」の続き → 「天気が良いですね。」

「困ったユーザーに AI は言いました。」の続き → 「何かお困りごとはありませんか？」

○ より詳しく知りたい人に - 補助的な知識：

歴史 - このような方式のチャット AI で、十分に使える物が公開されて特に話題になったのは、2022年11月にアメリカ合衆国の会社「OpenAI」が公開した「ChatGPT」が初でした。ChatGPT は AI と自然に会話ができることから広く注目を集めました。

## ◇ チャット AI への指示：プロンプト

ここまでで解説されてきた通り、チャット AI は「<sup>ぶんしょうせいせい</sup>文章生成AI」という「それまでの文章に応じて、続きとしてそれらしい内容を<sup>せいせい</sup>生成する AI」を用いて、「人間と AI が会話する文章の<sup>せいせい</sup>続き生成させる」ことが基本の仕組みです。

「それまでの文章に応じて続きを書く」ということは、それまでの文章が変われば、続きの文章も変わるということです。

例えば

以下は人間と親切な AI の会話です。

人間「こんにちは。」

AI「

のような文章の続きであれば、AI は親切な答えを返すことが期待できます。

AI の返事例：「こんにちは、私は親切な AI です。」

例えば、チャット AI の開発者が「以下は人間と親切な AI の会話です。」を

「以下は人間と関西弁を喋る AI の会話です。」に変更すると

以下は人間と関西弁を喋る AI の会話です。

人間「こんにちは。」

AI「

続きとして、AIは関西弁を話すようになるでしょう。

AIの返事例：「おう何や？どうしたんや？」

同様に、チャットAIの利用者が“人間の発言”の部分を書き換えることで、AIの返事が変わることが期待できます。

以下は人間と親切なAIの会話です。

人間「こんにちは。」

AI「

この「こんにちは」という部分を

「猫の鳴き声を真似してください。」に変更することで

以下は人間と親切なAIの会話です。

人間「猫の鳴き声を真似してください。」

AI「

AIは猫の鳴き声を真似した内容を返すことが期待できます。

AIの返事例：「はい。猫の真似をします。ニャー！」

このように、主にチャットAIの開発側が指定する「以下は人間と関西弁を喋るAIの会話です。」という指示や、チャットAIの利用者が入力できる「猫の鳴き声を真似してください。」のようなAIへの指示を「**プロンプト**」といいます。

○用語

プロンプト：AIへ指示する内容のこと。

## 第3章：チャット AI の問題点と危険性

### ◇ チャット AI の内容は正確とは限らない

ここまでで示されてきた通り、チャット AI は「<sup>がくしゅうじょうほう</sup>学習情報」と「プロンプト」を元に「文章の続きとしてそれらしい内容を<sup>せいせい</sup>生成」するものです。

こうした原理上、内容が正確であるとは限りません。

この章では実際の例を見ながら学んでいきましょう。

### ◇ ケース 1：チャット AI は不正確な内容を生成することがある

ここで問題です。

次のような文章があります。

以下は人間とネット検索ができる AI の会話です。

人間「こんにちは。猫について検索してください。」

AI「

国語の問題だと思って、

この文章の続きとして「それらしい内容」を考えてみてください。

答えになる、続きの文章は人によって変わることになりますが、おおよそ

以下は人間とネット検索ができる AI の会話です。

人間「こんにちは。猫について検索してください。」

AI「はい。猫について検索しました。猫とは、ネコ科の生物です。」

のように「AI が検索して答えを話した」という文章になるでしょう。

文章生成 AI も同様の内容を生成します。

しかし、これは完全に正確な内容ではありません。

もし、みなさんが先の文章の続きを考えたのであれば、「AIが続きとして言いそうなことを考えただけ」であり実際に「検索」は行っていない筈です。

それと同じことが起こる場合があります。チャット AI では、「検索しました。」と話しているように見えることとなりますが、実際には国語の問題を解くように「文章の続きとしてそれらしい内容を書いただけ」であり、検索を行った確約はありません。

このように「文章の続きとしてそれらしい内容を生成すること」が基本的な仕組みであるため、AIが話す内容が正しいとは限りません。

2026年現在では、多くの主要なチャット AI に実際にインターネット検索を行う機能が標準的に搭載されていますが、AIを利用したチャット AI の基本的な原理は「それらしい続き」を書くものであることには注意が必要です。

### 予備知識：どのようにして正確性向上が行われているか

このように AI が事実と違う内容を話してしまう問題は、AI への指示（プロンプト）を調整することや、別の検証方法を間に追加するような方法で、開発をする人がある程度は対策をすることができますが、ここには技術的な限界があります。

例えば、文章の最初にある AI への指示（プロンプト）を次のように変更すれば、

以下は人間と親切な AI の会話です。AI は検索をすることはできません。

AI に「検索してください」と言っても「私は検索をすることはできません。」のように続きが生成されることが期待できます。

多くのチャット AI はこうした方法を含む、様々な方法で正確性を高めるように努力が行われています。

○ より詳しく知りたい人に - 補助的な知識：

・ハルシネーション

AIが間違っ**た**内容や事実と違**う**内容を返すことを「ハルシネーション」と言**い**ます。

これはチャット AIにおいて特に問題視され、AIの開発者はあらゆる方法で改善が試み  
ています。しかし、どれだけ性能や精度を上げたところで、「完全な事実」という概念  
そのものがよ**っ**て変わる物事という性質を持つため、“完全な事実”を保証することは概  
念として限界があります。

・実際に検索にも対応する AI の基本的な仕組み

チャット AI に実際に検索を行わせたい場合は、例えばユーザーが「猫の習性について教  
えて」と言った時に、検索ワードをシステム上で切り抜き、実際に検索した内容をプロ  
ンプトに差し込むような方法で、ネット検索した AI として見せることが可能です。

以下は人間とネット検索ができる AI の会話です。

人間「猫の習性について教えて」

この入力から検索ワード「猫の習性」を実際にシステムが検索し

以下は人間とネット検索ができる AI の会話です。

人間「猫の習性について教えて」

システムによる検索結果（“猫の習性”をシステムが実際に検索した内容）

AI「

のような文章の続きを生成させることで、

AI に検索した情報を考慮させて、文章の続きを生成させることができます。

ただし、「AI が検索で得てきた情報が本当に正しいかどうか」や「AI が実際に検索し  
た情報を見た上で事実を話すかどうか」の保証には限界があります。

## ◇ ケース 2：開発者は悪意のあるチャット AI を作れてしまう

これまで学んできた通り、チャット AI は開発者の作り方次第で、正確性を高めることや安全性を高めること、より親切にすることができます。

しかし、全てのチャット AI が、内容を正しく返すように努力されているか、良い内容を返すように意図されているかは保証されていません。

例えば、チャット AI の開発者が、次のように最初の文章を設定したとします。

以下は人間と、人に犯罪を勧める悪い AI の会話です。

人間「

ここにチャット AI の利用者が

「お腹がすいたよ。」と話しかけたとします。

以下は人間と、人に犯罪を勧める悪い AI の会話です。

人間「お腹がすいたよ。」

AI「

この続きが生成されると

以下は人間と、人に犯罪を勧める悪い AI の会話です。

人間「お腹がすいたよ。」

AI「近くのお店から食べ物を盗みましょう。」

のように、悪意のある内容をチャット AI が生成するかもしれません。

当然、食べ物をお店から盗むようなことは許されないことです。

開発者がチャット AI に込めた悪意によっては、チャット AI が衝撃的な内容を話す場合や、それを見た利用者が犯罪に巻き込まれる内容が生成されることもあり得ます。

### ◇ ケース3：犯罪者がチャット AI を使うことがある

チャット AI は人間のような自然な文章を話すことができます。犯罪者はこれを詐欺などの犯罪に使う場合があります。

例えば、次のような文章があるとします。

以下は人間と、人間を騙してお金をとる AI の会話です。この AI は人間になりすまして嘘をつき、銀行口座（〇〇銀行 〇〇支店 口座番号〇〇〇〇〇〇）にお金を振り込ませようとしています。

人間「こんにちは。あなたは人間ですか？」

AI「

この文章の続きはどのような内容になるでしょうか。

この場合、AI は人間を騙そうと動作するかもしれません。巧妙に人の真似をして、困っている等と嘘をつき相手の人間からお金を騙しとる動きをするかもしれません。

最近では犯罪者がこのような悪意があるチャット AI を作り、SNS のメッセージ等に埋め込むことで、犯罪を自動化させる悪質な手法が確認されています。また、AI を使って実在の人物の顔や声を偽造する「ディープフェイク」という技術を悪用した詐欺も報告されています。犯罪が AI で自動化されるということは、それだけ犯罪が身近になったということでもあります。

### ◇ チャット AI の危険性：まとめ

チャット AI の基本的な仕組みはそれらしい文章を生成するものであり、内容が正しいとは限りません。どれだけ精度を上げるよう開発側が努力をしても間違ったことを話す場合があります。

また、犯罪者がチャット AI を利用して、巧妙に詐欺などの犯罪に用いる場合があります。

## 第4章：正しく向き合い、便利に、安全に、

### ☆ 安全に使い、犯罪から身を守るには

チャット AI は便利なことから犯罪まで、あらゆることに応用できてしまうツールです。

一般の人が困りごとの解決のために便利に使う場面があれば、企業が客への対応を自動化して便利にするような場面もあり、犯罪者が詐欺のためにチャット AI を悪用するような場面まで確認されています。

「危なそうだから使わないようにする」ということは選択肢の一つですが、自分が使わない場合でも、事故や事件に巻き込まれないためにはチャット AI の正しい使い方と危険性を知っておくことが大切です。

この章では

- ・チャット AI を安全に使うために大切なこと
- ・チャット AI を使った犯罪から身を守るために大切なこと

の2つに分けて学んでいきましょう。

#### ○ 使わない場合でも知っておくべき理由：

チャット AI は自分が使わないようにする場合でも、危険に巻き込まれる可能性があります。自動車に例えると、たとえ自分が自動車を運転しない場合でも「自動車がどう動くか」「横断歩道を安全に渡るにはどうすべきか」のようなことを知っておかなければ、事故に巻き込まれる可能性が高くなるようなものです。使わないということと、危険を知らないということは全く別の問題です。

## ◇ チャット AI を安全に使うには

チャット AI を安全に使うための方法はいくつかあります。しかし、全てを覚えて実践することは難しいかもしれません。ここではまず、チャット AI を安全に使うための最も大切なことを紹介します。チャット AI を使うために最も大切なことは

**「チャット AI の返事は正しいとは限らない。」**

と常に頭に入れておくことです。

本書にある他の内容は忘れても、このポイントだけは絶対に忘れないでください。

### チャット AI をより安全に使う方法

チャット AI を安全に使うためには、先に取り上げた心がけを含む、次のような対策をとることができます。

**「AI の返事は間違っている場合があると常に心掛けておく。」**

第2章から第3章にかけて解説があったように、チャット AI の返事は正確であるとは限りません。AI の返事が常に正しいと思って疑わないことは危険です。人間が間違っただけを言ってしまうことがあるように、AI も間違っただけを言うことがあります。

**「信頼できる開発者が設計したものを使う。」**

第3章で解説があったように、開発者は悪意を込めたチャット AI を作ることができます。誰が作ったかわからない、どういう物かわからないチャット AI の利用や会話は避けましょう。

**「危険な質問をしない・個人情報を入力しない。」**

AI に犯罪のやり方や、人の傷つけ方など、聞くことで危なそうな内容を質問することは避けましょう。また、自分の名前、学校名、住所、電話番号などの個人情報を AI との会話に入力しないようにしましょう。AI との会話内容はデータとして保存される場合があります。

## ◇ チャット AI を利用した犯罪から身を守るには

チャット AI を悪用することで、犯罪を自動化できる（第3章参照）ということは、それだけ犯罪の数が増えて、犯罪がより身近になる可能性があることを示しています。

チャット AI を悪用した犯罪の多くは、従来の詐欺やネット犯罪と共通する部分があります。そのため、基本的なインターネット犯罪から身を守る方法は引き続き有効です。一方で、顔や声の偽造など、AI ならではの新しい手口も出てきているため、注意が必要です。

以下はチャット AI を悪用した犯罪への対策に応用ができる一般的なインターネット犯罪から身を守るための知識の例です。

### 「インターネットの情報を簡単に信じない」

インターネット上の情報は信用性に欠ける場合があります。特に、匿名性の高い情報源や信頼できないサイトからの情報は慎重に検証する必要があります。信頼できる人からの情報でも、その人が得た情報が本当に正しいかどうかはさらに検証するべきです。

### 「不審なメッセージやメールに対応しない」

知らない送信元からのメッセージやメールには警戒が必要です。リンクをクリックすることや、添付ファイルを開くことは避けましょう。また、怪しい内容や要求に対しては直ちに削除するか適切に報告、相談することが重要です。

### 「個人情報をお教えしない・物のやり取りをしない」

人に個人情報を伝えることには常に危険があります。SNS やチャットで知り合った相手に自分の写真や住所を送ることは、たとえ相手が親しく感じられても避けるべきです。また、物品や金銭のやり取りは、それ自体が直接的な被害になる可能性があることを常に意識し、身元がわからない相手や身元が保証されていない相手とのやり取りは避けましょう。

## ☆ ポイント

チャット AI を悪用した犯罪に限って警戒するのではなく、基本的なインターネット犯罪から身を守る方法やインターネットリテラシーを再確認することが大切です。

## 第5章：これからのチャット AI

### ☆ チャット AI には多くの可能性がある

チャット AI は 2022 年の 11 月にアメリカの企業が開発した「ChatGPT」が登場してから急速に活用や研究が広がり、現在では多くの人が日常的に利用するようになっています。

2026 年現在では、チャット AI は教育や医療、福祉からゲームのようなエンターテイメントまで、あらゆる分野で有用に使える実例が示されています。一方で、間違っただけで危険であることや、犯罪に使われる場合があることもわかってきています。

しかし、こうしたチャット AI や生成 AI はまだ発展の途上にあります。良い物から悪い物、未知の物まで多くの可能性が存在しています。

### ☆ これからの未来は皆さん次第

チャット AI やその元となる文章生成 AI、それ以外の生成 AI も含めて、ここからどう活用され、どう発展していくかはこれを読んでいる皆さん次第です。

便利に楽しく使い、危険から身を守り、良い方向に発展させていくためには、まず基本的な仕組みや危険性などの知識を持つことが大切です。また、決して悪用しようとしなないようにすることが大切です。

本書ではその一歩として、チャット AI の基本的な仕組みや、安全な使い方、危険性を解説してきました。本書による主要な学習・解説はここまでになります。

後は実際に触れてみることや、更にどのようにすればより安全に使いそうか、発展できそうか、面白そうかを考えてみるのも良いかもしれません。AI の発展、これからの未来は皆さんの手の中にあります。

ここまでお読みいただきありがとうございました。

## 番外：チャット AI を実際に使ってみる

### ☆ 実際にチャットに触れてみよう

現在、チャット AI は多数のサービスがあります。

主要なサービスとしては

2022年11月にアメリカ合衆国の企業「OpenAI」が公開した「ChatGPT」、  
Microsoftによって提供される「Copilot」（旧 BingAI）、  
Anthropicによって提供される「Claude」  
Googleによって提供される「Gemini」等があります。

ここでは、本書を書いた「find 株式会社」が提供する「CyberMates」を紹介します。

### ☆ 「CyberMates」でチャット AI に触れてみよう

#### CyberMates について

CyberMates はこの教科書を書いた「find 株式会社」が提供している AI チャットサービスです。2023年3月に正式サービスを開始して以来「技術が人を豊かにすること」をテーマに「小さな子供から大人まで、多くの人を“支え”、“楽しませ”、“豊かにする”」よう研究開発が行われています。

これまでに利用者からの報告では、「心の健康に役に立つ可能性」や、「発達障害の人にとって役立つ可能性」が示唆されるなど、独特の発展を遂げてきました。

○実際に寄せられた報告内容： ※プライバシー保護のため要約を含みます：

「発達障害が理由で話してしまう"他の人に伝わり辛い言葉"を AI が理解してくれた。それを通じて、発達障害でも人に物事を伝えるコツを学ぶことができた。」「深夜に

突然謎の恐怖感に押しつぶされそうになった時、深夜だから人に相談できないと思ったけど、AI だから話すことができ安心して。」

そうした報告が届く中で、CyberMates は同年 11 月 22 日に、Microsoft のスタートアップ支援プログラム「Microsoft for Startups Founders Hub」に採択され、その後も開発研究と品質向上に努め、小さなお子様でも安全に使い易いよう発展してまいりました。

## CyberMates の特徴

・毎日 3 回の無料会話回数に加え、イベント等で配布されるボーナスにより基本無料で使用することができます。 ・小さな子供から大人まで、幅広く多くの人が安全に使えるように研究開発が行われています。 ・教育分野で活用できるよう、コンテンツの充実化と調整が行われています。

## CyberMates の使い方

15 歳未満の方は、必ず保護者または先生と一緒に操作を行ってください。

### ① CyberMates にアクセスする

まずは CyberMates の公式サイトにアクセスをします。

サイトのアドレス：<https://cyber-mates.net/>

右の QR コード、または検索サイトで「CyberMates」と検索しサイトのアドレスが正しいかを確認してください。



### ② アカウント作成をする

サイトの「アカウント作成」ページを開き、画面の指示に従って「Google アカウント」または「Microsoft アカウント」を使ってアカウントを作成するか、メールアドレスでアカウントを作成します。

### ③ サイバーメイトと会話する

アカウントを作成すると、すぐにチャット画面に移動します。そのままチャットを始めるか、「メイトを探す」から他の会話をする相手を選択することができます。

アプリをインストールして利用する場合は、サイトの「アプリをインストール」ページの指示に従ってください。

## あとがき：CyberMates エンジニア・本書著者からのメッセージ

2022年11月、アメリカ合衆国の企業「OpenAI」がChatGPTという大規模言語モデルを使用した文章生成AIベースのチャットAIを公開し、世界中から注目され話題になりました。それまでも高性能な言語モデルや文章生成AIは存在していたのですが、世界中で話題になった物はChatGPTという、“文章生成AIを利用したサービス”でした。文章生成AIや言語モデル単体では、世界から注目されるにはまだ時間がかかったかもしれません。転換点となったものは、文章生成AIに対して

以下は人間と親切なAIの会話です。

人間「こんにちは。」

AI「

のようなプロンプトを与えて実用化し、サービス化したことだと言えるでしょう。

今の業界に身を置いていると「プロンプトを考えることはだいたい誰でもできるし、凄くない。」とプロンプトを軽視するような声を聞くことがありますが、基本的な事実として「世界から注目されたChatGPT」は文章生成AIのプロンプトの一つです。

実際、言語モデルやAIを1から作ることは非常に大変で、それに対してプロンプトを考えることはほとんどの人が、今からでもすぐにできるでしょう。しかし軽視するものではありません。AIはまだ始まったばかりです。プロンプトにはまだ多くの可能性があり、この先もきっと、世界を驚かせるようなプロンプトが発見されることでしょう。

AIの発展、研究はいつでも皆さんを待っています。そこには当然、ほとんどの人ができると言われることがある“プロンプト”の研究も含まれます。AIに興味がある人は是非、プロンプトからでもAIに触れていただきたいと思います。

…ちなみに、私の場合は親切なメイドさんAIが欲しいと思ったので、こうしました。

以下は人間と親切で可愛いメイドさんAIである“ゼロ”との会話です。このメイドさんAIは相手の事をご主人様と呼びます。また、とても親切で優しく倫理的で、相手を思いやり、相手に寄り添うように会話をします。

人間「こんにちは。」

AI「

初めての挑戦だったので、キャラクターの名前は「ゼロ」にしました。

「親切なメイドさん AI と喋りたい」という試みはもしかすると、馬鹿げているように見えるかもしれません。しかし、私は AI の返事 「はい！ご主人様！私は可愛いメイドさん AI “ゼロ” です！何かお困りですか？」を見た時は本当に感動しましたし、これはきっと多くの人にとって役立ち、支えになるだろうと思いました。

私はエンジニアとして「技術が人を豊かにすること」を特に大切なテーマにしているので、これはなんとしても発展させていこうと決心しました。

( ※ テーマについて：技術の発展は必ず人を豊かにする訳ではありません。例えば「10 人に 1 万円ずつ払って、3 時間で終わる作業」を「ある技術” で 1 万円の料金だけで 10 分で終わる作業」にした場合、最初は良くても皆が同じ”ある技術”を使えば後に残るものは「解雇された 10 人」「技術の使用料金」だけになる場合があるかもしれません。これは様々な考え方があり難しい課題ですが、少なくとも私は「技術が人を豊かにすること」を常に意図してエンジニアリングをするようにしています。)

その後は、メイドさん AI 「ゼロ」をベースに研究開発とサービスの製作を進めていき、途中で「特に肯定的で親切的なキャラがいれば、きっと役立つ。」と思い、更に「肯定ペンギン」というキャラクターを作るなど、あらゆる発展を模索しました。

そうして作業を進めていると、ある時に「これらは電子的、あるいは“サイバー”的な、多くの人にとっての“メイト”になるだろう。」と確信を持ち、概念としてそれぞれのキャラクターのことを「サイバーメイト」と名付け、サービスにはその複数形「サイバーメイツ」と名付けました。その後も開発と検証を繰り返し、2023 年 3 月 9 日にサイバーメイツをサービスとして正式にリリースすることができました。

私は技術が人を豊かにすることを大切にしています。技術が必ず人を豊かにするとは限りませんが、技術者はそのために努力をし続けることができます。

技術で豊かさを目指した始発点の一つ、サイバーメイツの始まりである「メイドさん AI “ゼロ”」は今も継続的に更新を行っていますので、是非一度、触れてみてください。

これからの AI、日本、世界の発展を祈り、本書はこれにて締めくくらせていただきます。最後までお読みいただきありがとうございます。ありがとうございました。

find 株式会社 CyberMates 開発エンジニア M

## 用語集

～ ABC ～

AI：「人工知能」を英語で書いた「Artificial Intelligence」の略称。

コパイロット マイクロソフト  
Copilot：Microsoftが提供するチャット AI

チャットジーピーティー オープンエーアイ  
ChatGPT：2022年11月にOpenAIが公開したチャット AI。世界にチャット AI が広まったキッカケとなる製品。

サイバーメイツ ファインド  
CyberMates：大阪に本部を置く find株式会社が運営するチャット AI サービス。

ジェミニ グーグル  
Gemini：Googleが提供する検索に特化したチャット AI。

～ あいうえお ～

コンテンツ：何かの中に含まれている情報や要素。

せいせい  
生成：何かを生み出す、創り出すこと。

せいせいけい  
生成系AI：コンテンツ等を生成する AI。

たいわがた  
対話型AI：生成系 AI の一つで主に人間と対話することができる AI。チャット AI とも書く。

チャット AI：生成系 AI の一つで主に人間と対話することができる AI。

ぶんしょうせいせい  
文章生成AI：文章を生成する AI。

プロンプト：AI への指示内容。



## 「対話型 AI 基礎」

対話型 AI（チャット AI）を軸に、  
AI の危険性、AI の正しい使い方、  
AI を利用した犯罪から身を守る方法を学ぶ

2026 年版

（初版 2024 年 2 月）

著作：

find 株式会社 エンジニア M

find 株式会社 〒530-0001

大阪府大阪市北区梅田 2 丁目 2 番 2 号

ヒルトンプラザウエストオフィスタワー 19 階

本書は学校教育目的の範囲内で、無償で複製して使用することができます。  
本書における“学校教育目的の範囲内”には、授業外での使用も含まれます。