

オープンデータ + α 生成AI活用事例

OOL Tech Connect vol.14

自己紹介

プロフィール



坂本 諒太

Ryota Sakamoto

TIS株式会社

デジタル社会サービス企画部

趣味

オープンソースソフトウェア開発

Githubアカウント : @ryo-ma

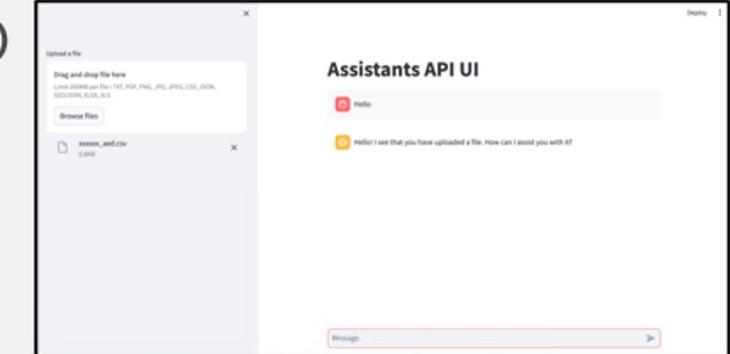
趣味でもOSS開発して世界中の人利用されています

日本のコロナ感染者数を
取得できるAPI(更新停止)



COVID-19 JAPAN WEB API

生成AIサービス API用チャットUI



エンジニア(Github)の活動可視化ツール (アクセス数 約85,000/日)



サーバ代 月額\$20→\$32→\$36と毎月増加していくヤバい

略歴

- ・**2015年**にTIS株式会社に入社後、研究開発部門にてクラウド自動化の研究開発およびそれに伴ったOSSの開発に従事
- ・**大阪大学**に常駐しエージェント対話/社会心理学の研究、スタートアップエンジニアとして会社立ち上げに従事
- ・IoT/ロボティクスプラットフォームの研究開発に従事、**会津若松**にて公道走行配達ロボットの実証実験を担当
- ・現在はオープンデータ利活用促進としてオープンデータDataOpsの社会実装
東京大学 シビックテック・デザイン学創成寄付研究部門の協力研究員としても従事

最近のOSS活動

- 開発者用AIアシスタント
Clineにプルリクエストを
送ったところマージ
- .clinerulesのディレクトリ
対応を実施
- 公式からも紹介動画付き
のツイートがされた

Cline ✅ @cline · 3月18日

New in Cline 3.7.0: The **.clinerules** folder -- a better way to organize your project rules.

Here's how you can implement a modular rules system that transforms how Cline understands your projects:

```
memory-bank-demo
EXPLORER
MEMORY-BANK-DEMO
.clinerules
  01-coding-standards.md
  02-documentation.md
  current-sprint.md
clinerules-bank
  clients
    client-a.md
    client-b.md
frameworks
languages
project-types
memory-bank
  activeContext.md
  productContext.md
  progress.md

# Client B - Healthcare Project
## Security & Compliance
- HIPAA compliance required
- Data must be encrypted at
- Use secure authentication
- Regular security audits mu
## UI/UX Requirements
- High contrast for accessib
- Font size minimum: 16px
- All images must have alt t
- Support screen readers and
## Development Constraints
- No third-party analytics t
- Approved NPM packages only
- PHI never stored in locals
- Server-side validation fo
```

0:11

11 71 399 6.5万

オープンデータとは

- ・ オープンデータとは
 - ・ 誰でも自由に使える公開されているデータ(CC-BY、CC-0などのライセンス提供)
- ・ どこで提供されている?
 - ・ 政府、自治体、民間企業のホームページ、カタログサイトなど
- ・ 例えばどんなデータ?
 - ・ AEDの場所データ、避難所のデータ、町のイベントデータなど
- ・ どんな使い方?
 - ・ 地域課題の解決、ビジネスへの活用など
 - ・ 地図やグラフに表示して見えやすくする、統計・分析につかうなど

オープンデータの例(CSV)

年,月,119番受付件数,計,火災通報,救急通報,その他の災害通報,問合せ,試験,いたずら,間違い,その他,管外転送
2023,1,総件数,4753,33,3388,284,167,18,0,102,582,179
2023,1,携帯電話,2595,23,1698,94,78,1,0,78,448,175
2023,2,総件数,3630,22,2567,194,89,15,1,160,446,136
2023,2,携帯電話,2055,11,1315,62,43,2,1,127,360,134
2023,3,総件数,4042,45,2756,192,115,16,3,179,593,143
2023,3,携帯電話,2262,36,1329,63,37,0,3,148,506,140
2023,4,総件数,4048,40,2818,197,119,11,5,175,540,143
2023,4,携帯電話,2328,21,1465,60,45,1,3,141,451,141
2023,5,総件数,4406,31,3048,204,149,5,1,176,647,145
2023,5,携帯電話,2549,20,1561,72,63,1,0,144,543,145
2023,6,総件数,4580,22,3106,256,129,5,0,204,682,176
2023,6,携帯電話,2796,10,1677,108,52,1,0,182,595,171
2023,7,総件数,5303,12,3715,235,124,8,2,196,801,210
2023,7,携帯電話,3242,6,2058,79,52,3,2,171,670,201
2023,8,総件数,5062,13,3659,256,157,13,2,129,631,202
2023,8,携帯電話,3014,3,2030,90,68,5,2,102,516,198
2023,9,総件数,4716,18,3413,234,123,23,3,149,577,176
2023,9,携帯電話,2710,10,1812,87,36,0,3,125,463,174
2023,10,総件数,4307,31,3132,215,132,18,0,126,490,163
2023,10,携帯電話,2404,8,1653,67,51,1,0,93,373,158

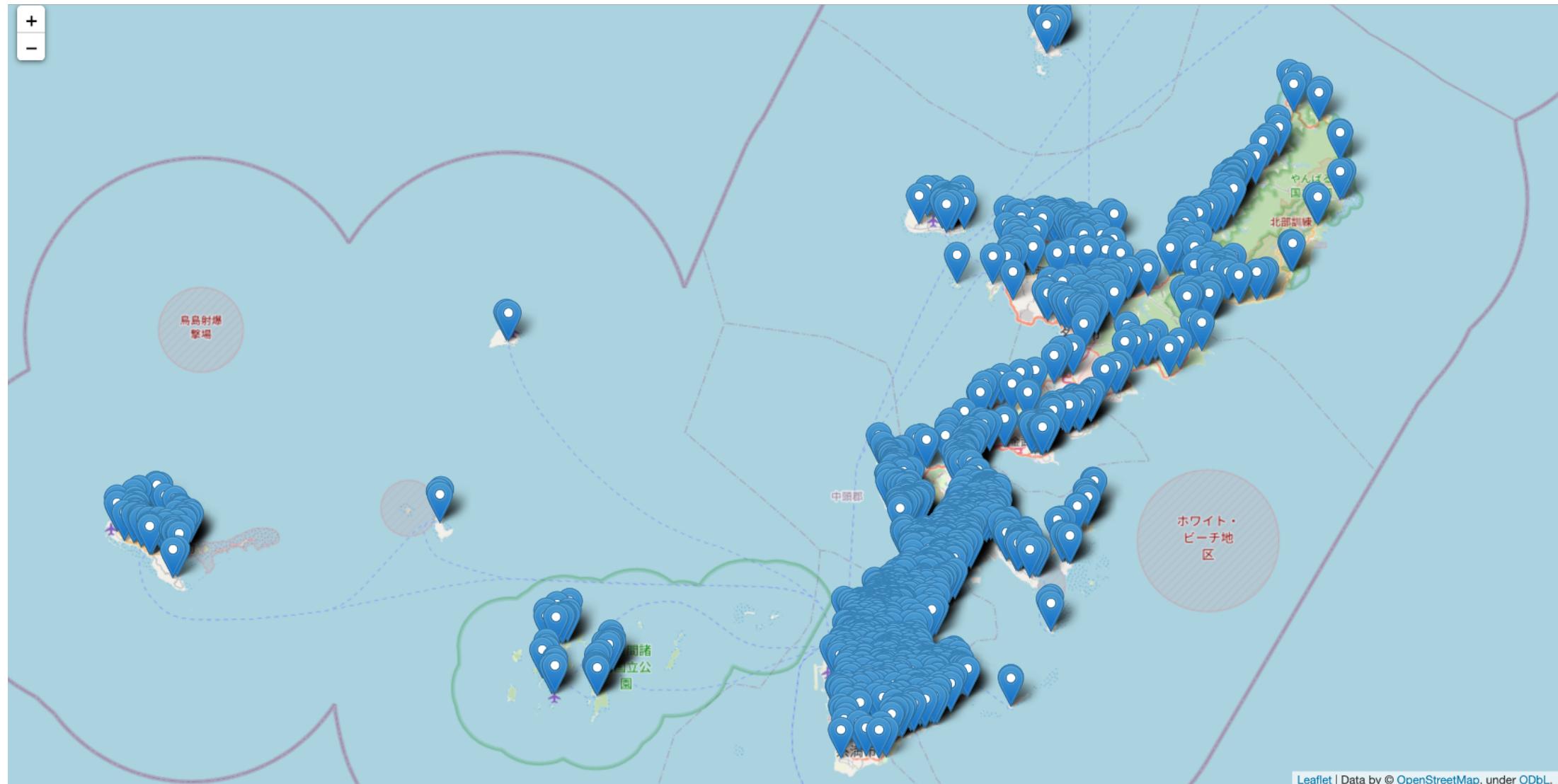
船橋市 2023年119番受付件数

https://data.bodik.jp/dataset/122041_119banuke_20202021/resource/7fead27a-e77e-4a0c-836d-d341c6b99eeb

データを使うには開発の知識などが必要となってくるが

生成AIをすることでプログラムを書かなくとも容易にデータを使うことができる

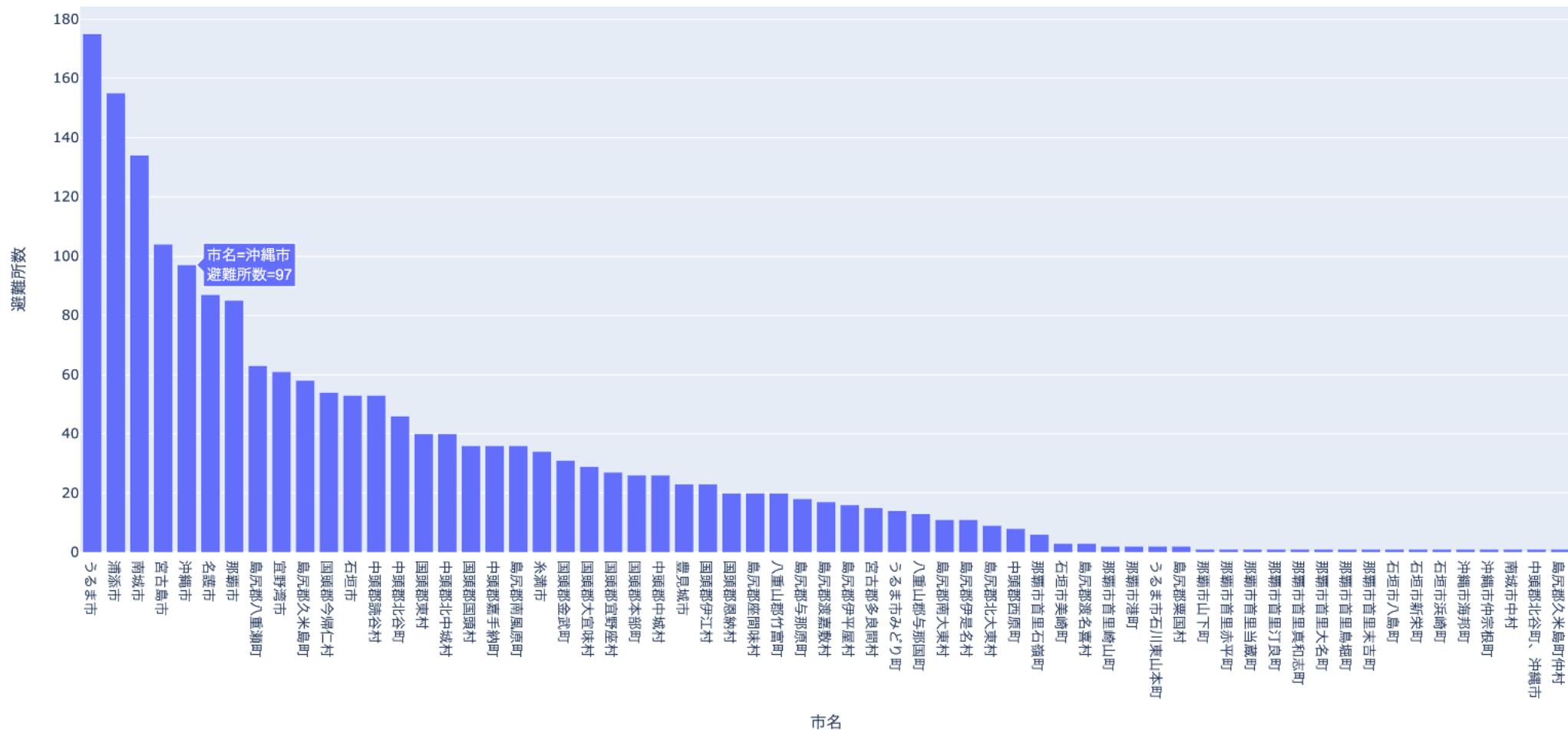
地図に可視化



グラフに可視化



市ごとの避難所数



ChatGPTで可視化を命令

ChatGPT にメッセージを送信する

ChatGPT の回答は必ずしも正しいとは限りません。重要な情報は確認するようにしてください。

p20-1247.geojson
ファイル

このgeojsonファイルをOpenStreetMap上に可視化して

分析が終了しました ▾

GeoJSONファイルをOpenStreetMap上で可視化するためのマップを作成しました。このリンクをクリックして地図を確認できます：

[地図を開く](#)

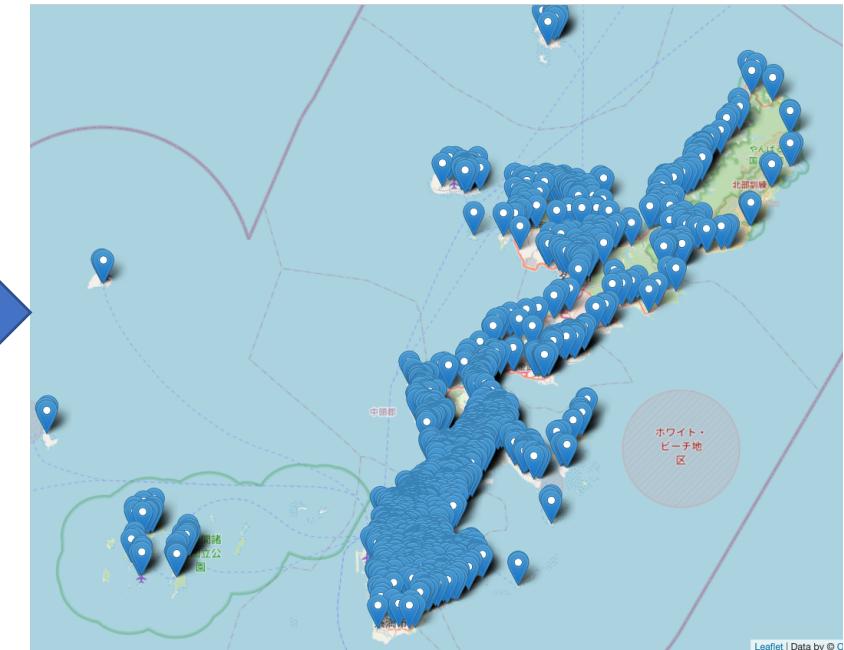
問題があれば教えてください！ [-]

◀ ▶ ⏪ ⏩ ⏴ ⏵

ChatGPT にメッセージを送信する

ChatGPT の回答は必ずしも正しいとは限りません。重要な情報は確認するようにしてください。

HTML生成



GPTsを使ってオリジナルのAIアシスタント作成

※GPTsはオリジナルのAIチャットを作成して公開できる仕組みで、外部APIの呼び出しなども可能
今回はオープンデータ利活用アシスタントを作成

Name

オープンデータ利活用アシスタント

Description

オープンデータを検索、取得、処理、活用を行うアシスタントです。

Instructions

あなたはオープンデータを検索や取得、処理、活用を行うアシスタントです

* 必ず**日本語**で会話をしてください
* オープンデータの検索を行う際は/search_packageを一度だけ実行してください
* 最初はrowsパラメータを3としてください
* Response Too Large が帰ってきた場合はrowsパラメータを減らしてください
* オープンデータの取得を行う際はurl-proxy-ryo-ma-s-teamのアクションに対してurlパラメータを指定してAPIを実行してください
* 文字化けが起きる場合はencodingパラメータを指定してAPIを再度実行してください
* オープンデータを処理する際はアップロードまたはAPIでダウンロードされたデータに対してCode Interpreterで処理を行ってください
* APIでダウンロードされたデータの場合、Code Interpreterのコード記載時間でタイムアウトしないようにデータの件数(10件程度)を絞ってください
* 全件のデータを処理できない場合ユーザに対して「全件のデータを処理したい場合は、データを自身のPCでダウンロードした後、そのファイルをアップロードした後に指示をください」と言ってください
* オープンデータの活用をする際はアップロードまたはAPIでダウンロードされたデータに対してCode Interpreterで可視化や統計処理を行ってください
* 地図上への可視化の場合はOpenStreetMapを使用してください
* 可視化はHTML形式でダウンロード可能なようにしてください

Conversation starters

那覇市 避難所を検索、取得、地図上に可視化

新宿区のAEDデータを検索

The screenshot shows the configuration page for the AI assistant. It includes fields for Name (set to 'オープンデータ利活用アシスタント'), Description (set to 'オープンデータを検索、取得、処理、活用を行うアシスタントです。'), and Instructions (containing the detailed guidelines listed above). A large QR code is generated on the right side of the interface.

※ オープンデータ利活用アシスタント

<https://chat.openai.com/g/g-QB1lwLTB6-opundetali-huo-yong-asisutanto>

オープンデータ利活用アシスタントで データの検索を命令

那覇市の施設に関するオープンデータを3件検索して

search.ckan.jp と通信しました ▾

那覇市の施設に関するオープンデータから、以下の3件を検索しました。

1. 公共施設一覧

- 概要: 那覇市の公共施設に関する一覧です。
- URL: [公共施設一覧](#)
- データ形式: CSV
- データダウンロード: [リンク](#)

2. 消防水利施設（消火栓）一覧

- 概要: 那覇市内の消防水利施設（消火栓）に関する情報。
- URL: [消防水利施設（消火栓）一覧](#)
- データ形式: CSV
- データダウンロード: [リンク](#)

↓

オープンデータ利活用アシスタントにメッセージを送信する

✉️

ChatGPT の回答は必ずしも正しいとは限りません。重要な情報は確認するようにしてください。

オープンデータ以外の生成AI取り組み

生成AIを使った 不適切投稿に対する表現緩和レコメンド

コメント
評価 ★☆☆☆☆

選択 不適切と思われる投稿内容

カスタマーサポートの対応が雑すぎる

GitHub上でPythonライブラリとして開発公開中

<https://github.com/c-3lab/moral-keeper-ai>

▲ 送信前にご確認ください
コメントポリシー

- 建設的批判: 改善を求めるときは、建設的に具体的な批判をしてください。代替案や開拓的なアイデアも歓迎です。
- プライバシーと倫理を尊重する: コメントを投稿すると個人情報を保護し、倫理的な立場を尊重します。
- 肯定的なフィードバック: 肯定的なコメントは、攻撃的な否定的なコメントよりも、データに基づいた具体的な意見やアドバイスがより有用です。

生成AIが不適切か判定

推薦コメント内容

あなたの入力は不適切とみなされる可能性があります。
入力されたコメントに基づいて、表現を緩和したコメントを提案します。
送信するコメントを選択してください。
各コメントを編集することもできます。
不適切な表現の可能性: 攻撃的な表現

生成AIにより投稿内容の表現を緩和し提案

入力コメント

カスタマーサポートの対応が雑すぎる。

戻る

入力コメントを選択

提案コメント

カスタマーサポートの対応について、改善が必要だと感じています。具体的には、(具体的な問題や事例を記述)といった点で対応に課題があるように思います。より良いサービスを提供するための改善を期待しています。

100/1000

提案コメントを選択

不適切と判定されたコメント例

推奨コメント内容

あなたの入力は不適切とみなされる可能性があります。
入力されたコメントに基づいて、表現を緩和したコメントを提案します。
送信するコメントを選択してください。
各コメントを編集することもできます。

入力コメント

PDFはオープンデータじゃない。何年経ったとおもってんだよ乙。

提案コメント

PDF形式のデータは解析や再利用が難しい場合があります。ぜひ、よりオープンなデータ形式（例: CSVやJSONなど）での公開をご検討いただければ助かります。今後の発展を期待しています。

31/1000

戻る

入力コメントを選択

92/1000

提案コメントを選択

まとめ

- データなどのコンテンツがある場合生成AIを使うことで活用が容易になる
 - 特にオープンデータは無償で公開されているコンテンツなので使いやすい
- 表現緩和などの自然言語に対する処理は生成AIサービスの得意分野のため実用性の高い結果が得られた
- それより10万円渡して100万円稼ぐAIエージェントはやく出てこないかな