

OOL Tech Connect vol.4

生成AIを活用したオープンデータ利活用促進

2023.12.6

TIS株式会社

デジタル社会サービス企画ユニット

デジタル社会サービス企画部

坂本 諒太

December 06 – DAY2

09:00~12:00 (3h)

講演+ハンズオン

OOL TECH CONNECT

※ハンズオンに参加される方は別途Connpassからお申し込みが必要です。
(講演のみ聴講の場合はOOD登録でOK)
(※現地開催のみ)



13:30~14:15 (45min)

Keynote

光ファイバセンシングによる新たな価値の創造

青野 義明

日本電気株式会社 トランスポートネットワーク統括部 ディレクター



14:30~15:20

Session

14:30-14:50 (20min)

地域の持続、活性化に向けた取組みと未来社会に向けた投資

坂本 勇人 (伊藤忠テクノソリューションズ株式会社)

14:50-15:10 (20min)

共創による未来社会デザインと人的資本強化への挑戦

鎌滝 優紀子 (日本電気株式会社)

15:10-15:20 (10min)

Firewall・ネットワークテストプロセスにおけるデジタルイノベーション

國森 修 (APコミュニケーションズ)



December 07 – DAY3

09:30~11:00 (90min)

Panel Discussion

Trustable Supply Chain Frameworkの実現に向けて

- Chang-Kyu Kim (国立研究開発法人 産業技術総合研究所 サイバーフィジカルセキュリティ研究センター協力研究員)
- 伊賀 洋一 (一般財団法人日本規格協会)
- Allan Friedman (米国サイバーセキュリティ・社会基盤安全保障庁 (CISA))
- Michael Pease (米国国立標準技術研究所 (NIST))
- SZLin (林上智) (Chief Cybersecurity Expert, Bureau Veritas Taiwan/ President of the International Society of Automation Taiwan Section)
- 辻川 公章 (アラクスネットワークス株式会社)



11:15~12:15 (60min)

Session

OOL研究発表

Model Driven Network DevOpsプロジェクト

ISPネットワークのモデルベース再現とBGP運用シミュレーション

前野 洋史 (ビッグロブ株式会社)

山口 大樹 (伊藤忠テクノソリューションズ株式会社)

DataOpsプロジェクト

生成AI x DataOps 生成AIによるオープンデータ活用促進

宮岡 時生 (TIS株式会社)

Edgeプロジェクト

沖縄におけるMECの活用と地域データ利活用・連携の可能性

亀井 貴行 (NTTコミュニケーションズ株式会社)



12/7(金) 11:15-12:15

OOL研究発表にて
**「生成AI x DataOps 生成AIによる
 オープンデータ利活用促進」**
 にて本講演の詳細な内容を紹介します

1. 自己紹介
2. オープンデータにおける背景
3. オープンデータの課題を生成AIを使って解決、改善
4. 生成AIによる「オープンデータ作成」
5. 生成AIによる「オープンデータ活用」
6. 全体のまとめ



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

趣味など
OSS開発

SNS

Twitter
@SkeinIn

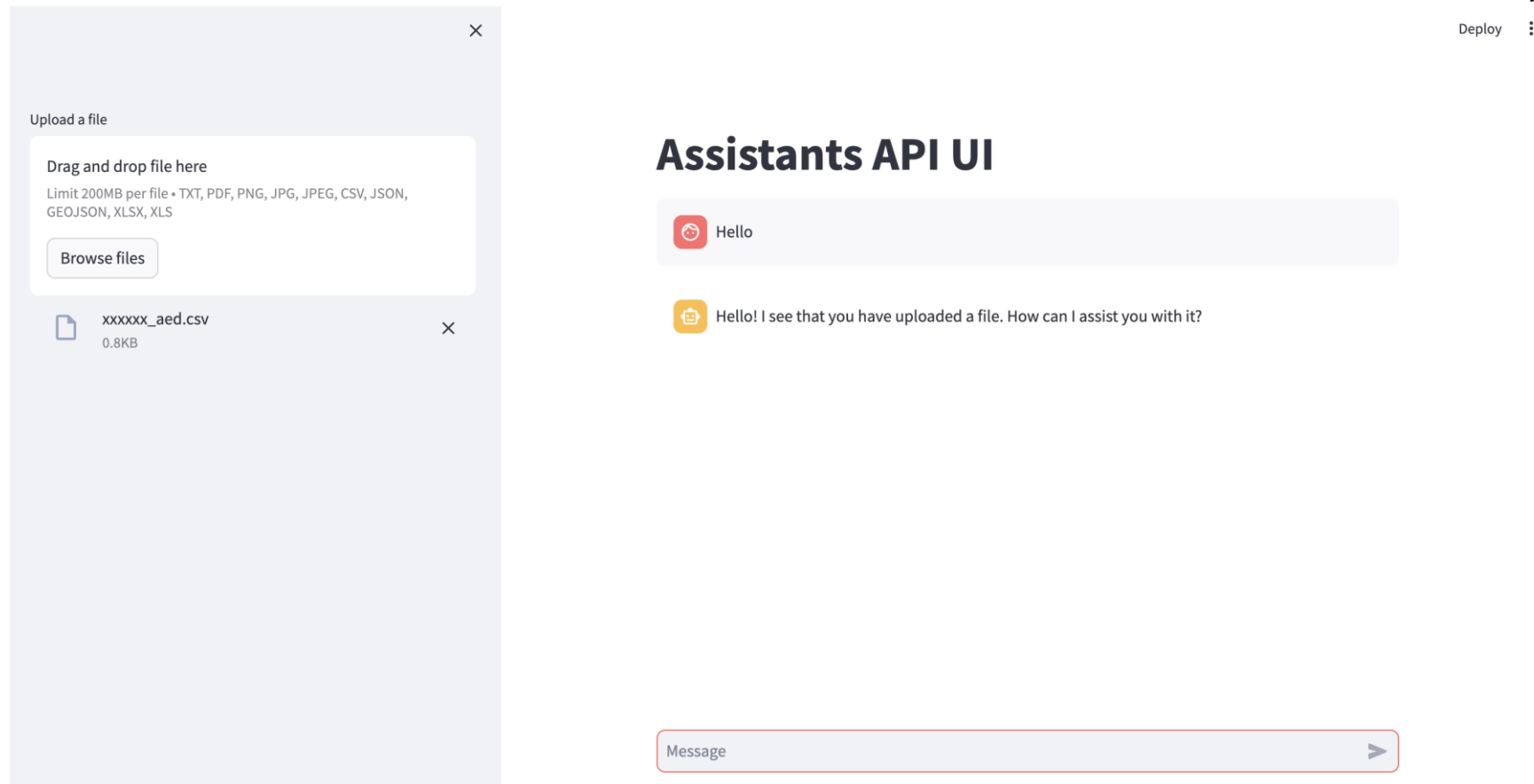
Github
@ryo-ma

略歴

- ・ 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
- ・ 大阪大学に常駐し対話/社会心理学の研究に携わり、その知見を活用した高齢者向け生活支援AI対話サービスをスタートアップエンジニアとしてスタートアップ立ち上げに従事
- ・ FIWAREによるIoT/ロボティクスプラットフォームの研究開発に従事し、会津若松にて公道走行配送ロボットの実証実験を担当
- ・ 現在はデータ利活用という観点で社会のあるべき姿を検討し社会実装に向けた活動を実施

OpenAIから発表された Assistants APIを実行できるUIをOSSとして公開

※OpenAI Assistants API：オリジナルの高度機能を持ったAIChatを作成できるAPI、ファイルのアップロードしファイル読み込み可能、ファイル生成しダウンロード可能、Pythonの実行に対応など

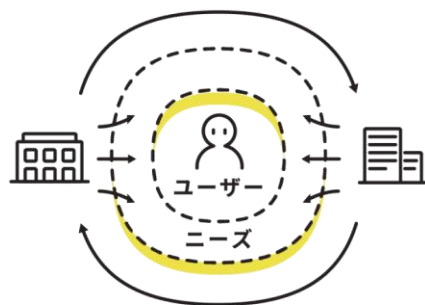


イベントなどでChatGPTアカウントを作らずに高度な機能付きAIChatを提供したりなどに活用

オープンデータにおける背景

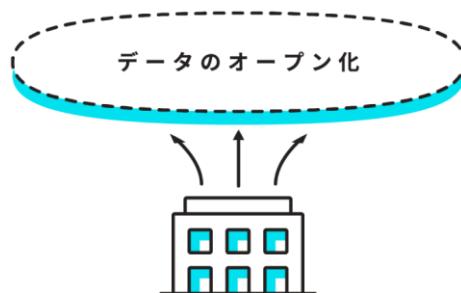
ITがあれば、全ての地域課題が解決されるわけではない。しかし自分たちの課題を自分たちで解決していく上で、ITが役に立つ場面も多い。

その中でも、その地域や周辺地域の情報に容易にアクセスできれば、様々な側面から地域課題の解決に役立つITシステムが形作られるのではないかと考えられており、オープンデータの活用は我が国の成長戦略の重要なパーツとして認識されている。※)



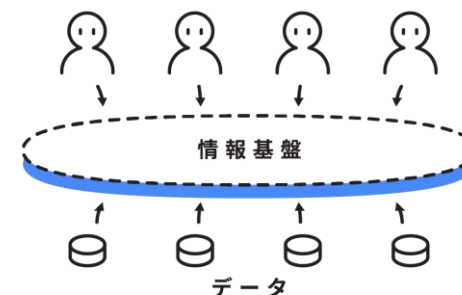
利用者視点に立つ

徹底的に利用者の視点に立ち、優れたサービス体験を実現するために、民間サービスとの連携も含めてシステムやルールを設計します。



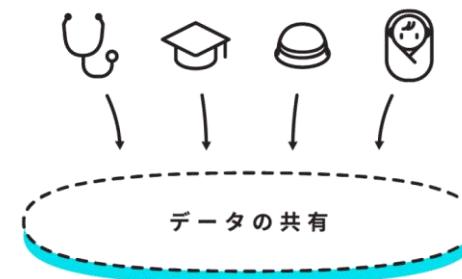
データの利活用を促進

様々な分野のデータを官民を越えて連動させることが、国民の多様な幸せの実現に不可欠です。そのためのデータのオープン化（一般公開）と環境の整備に取り組みます。



データを誰でも扱いやすく

国民が自らのデータを必要なときに素早く利用できるようにするため、また、府省間の連携を強めるための情報基盤を構築します。情報基盤は、扱いやすく標準化された形式・方法で活用できるようにします。



連携の仕組みの標準化

現在、各サービスはそれぞれの事情に合わせて個別に設計されています。データの標準化やルール整備により、官民や分野を横断する連携を実現します。

- **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
- **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

オープンデータの課題感は？

オープンデータの提供者側も利用者側も、オープンデータを活用したいと切に願っている。
しかしお互いの思いが直接伝わらないため、「とりあえずできるところから始めてみる」の先へ進むことができていない。

データ提供者（自治体や地場の企業等）



データ利用者（住民やシビックハッカー、企業等）



オープンデータの課題を生成AIを使って解決、改善

オープンデータの作成と活用における課題では生成AIを使うことで解決、改善につながると考えた

データ提供者の課題

オープンデータの作成

- データを綺麗に作れる
- データの間違いを知る
- 間違ったデータを修正する

生成AI活用



ChatGPT 等

データ利用者の課題

オープンデータの活用

- データを使えるように加工
- 地図にデータを出す
- グラフにデータを出す

1. オープンデータの知識がなくても作れる

2. 開発知識がなくてもデータが活用できるようになる

生成AIによる「オープンデータ作成」

世の中のオープンデータは使いにくい形式(PDF、XLSX)だったり
必要な項目が不足していたりする



正しいオープンデータの作り方を知る必要がある



オープンデータのレビューAIを作る



OpenAI社が公開したPythonの実行とアップロード・ダウンロードが可能なモデル
現在はChatGPT Plusでベータ版として提供が開始されている

◆ 特徴

- 外部環境から切り離されたサンドボックス上でPythonコードを実行可能
- 実行したコードはチャットを消さない限り、コードの再利用が可能
- チャットへのファイルアップロードが可能
- 実行結果のダウンロードが可能

◆ ユースケース

- 数学的な問題の解決 (定量的 / 定性的)
- データの分析と可視化
- ファイルのフォーマット変換

◆ 安全性について

- ネットワークの制限をかけることで現実世界へ予期せぬ作用をもたらさないように配慮
- それぞれのセッションのリソースを一定に制限

1. オープンデータのルールを
事前の知識として教える



ChatGPT
Code Interpreter

2. オープンデータのルールに
基づいて回答

命令：添付の資料を理解したら「はい」と言ってください

項目	項目名	属性	データ項目 (AED設置場所) 説明	形式	読み出し	関連項目	関連項目の属性	関連項目の項目名	関連項目の属性	関連項目の項目名
1	施設名称	文字列	施設名称	文字列	施設名称	施設名称	施設名称	施設名称	施設名称	施設名称
2	所在地	文字列	所在地	文字列	所在地	所在地	所在地	所在地	所在地	所在地
3	経度	文字列	経度	文字列	経度	経度	経度	経度	経度	経度
4	緯度	文字列	緯度	文字列	緯度	緯度	緯度	緯度	緯度	緯度
5	電話番号	文字列	電話番号	文字列	電話番号	電話番号	電話番号	電話番号	電話番号	電話番号
6	URL	文字列	URL	文字列	URL	URL	URL	URL	URL	URL
7	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
8	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
9	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
10	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
11	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
12	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
13	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
14	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
15	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ
16	施設タイプ	文字列	施設タイプ	文字列	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ	施設タイプ

デジタル庁 自治体標準オープンデータセットデータ項目定義書（正式版）
https://www.digital.go.jp/resources/open_data/municipal-standard-data-set-test/

命令：先ほど共有したデータ項目定義書に基づいてこの添付ファイルを
フィードバックしてください

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	施設名称	所在地	緯度	経度	電話番号	URL								
2	両国観光	墨田区両国	35.69547	139.7922	03-3631-1	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
3	両国小学	墨田区両国	35.69377	139.7952	03-3634-7	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
4	緑小学校	墨田区緑	35.69403	139.8013	03-3634-6	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
5	緑図書館	墨田区緑	35.69576	139.8006	03-3631-4	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
6	みどりこ	墨田区緑	35.69366	139.8033	03-5600-5	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
7	すみだ隣	墨田区緑	35.69551	139.8076	03-5600-2	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
8	すみだふ	墨田区緑	35.69582	139.8074	03-5600-2	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
9	立川児童	墨田区立川	35.69051	139.7995	03-3633-2	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
10	菊川小学	墨田区立川	35.69163	139.8076	03-3634-8	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
11	中和小学	墨田区菊川	35.69016	139.8027	03-3634-7	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
12	江東橋児	墨田区江東	35.69592	139.8097	03-3634-8	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
13	大横川親	墨田区江東	35.69397	139.8089	03-3633-5	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
14	すみだ産	墨田区江東	35.69505	139.8137	03-3635-4	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
15	弓道場	墨田区江東	35.69354	139.8193	03-3633-5	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								
16	墨田区水	墨田区江東	35.69266	139.8108	03-3635-5	http://www.city.sumida.lg.jp/kenko_fukushi/kenko/kokoro/settisisetu/aedsetup.html								

東京都墨田区 AEDオープンデータ
<https://www.geospatial.jp/ckan/dataset/13107-001>

2. オープンデータのルールに基づいて回答

修正が必要な点に加え、修正が必要な理由を出力してくれた。

※実際はChatGPTといくつかのやり取りが発生



ChatGPT
Code Interpreter

ファイル 文字コード変換
文字コードをshift-jis -> utf-8

回答

- CSV ファイルに「全国地方公共団体コード」および「ID」の列を追加してください。これにより、データの出所と一意性が明確になります。
- CSV ファイルに施設の名称のカナ表記と英字表記の列を追加してください。これにより、日本語が母国語でない人々や、カナ表記を必要とするシステムでもデータが使いやすくなります。
- 所在地情報を詳細化してください。特に、「都道府県」、「市区町村」、「町字」、「番地以下」を個別の列として提供することで、より詳細な地理的分析が可能となります。
- 「08.AED設置箇所一覧」シートに定義されている他の項目（例えば、「名称_カナ」、「名称_英字」、「所在地_全国地方公共団体コード」、「町字ID」など）についても、可能であればこれらの情報を提供してください。これにより、データの利用可能性と相互運用性が向上します。

新たな項目の提案が多かったので、既存の項目に対するフィードバックを追加で要求

※実際はChatGPTといくつかのやり取りが発生



ChatGPT

Code Interpreter

命令：既に記載されているデータに関して、データ項目定義書に則ったフィードバックはありますか？

回答

- 施設名称（名称）：既存のデータは定義書に準拠しています。特に改善点は見られません。
- 所在地: 定義書では所在地情報は「都道府県」、「市区町村」、「町字」、「番地以下」のように細かく分けられていますが、提供されたデータではこれらが一つの列にまとめられています。これらを個別の列に分けることで、データの活用性が高まります。
- 緯度・経度: これらのデータは定義書に準拠しています。ただし、測地系（例：世界測地系（WGS）や日本測地系）が明示されていない場合、これらの情報を明記することをお勧めします。
- 電話番号: 定義書には電話番号の項目は見られません。しかし、この情報は施設への連絡方法として重要なため、現状維持でも問題ないと思われます。
- URL: 定義書にはURLの項目は見られませんが、施設やAEDの詳細情報を提供する上で有用な情報であるため、現状維持でも問題ないと思われます。ただし、これらがリンク切れでないこと、または最新の情報を反映していることを定期的を確認することをお勧めします。

生成AIによる「オープンデータ活用」

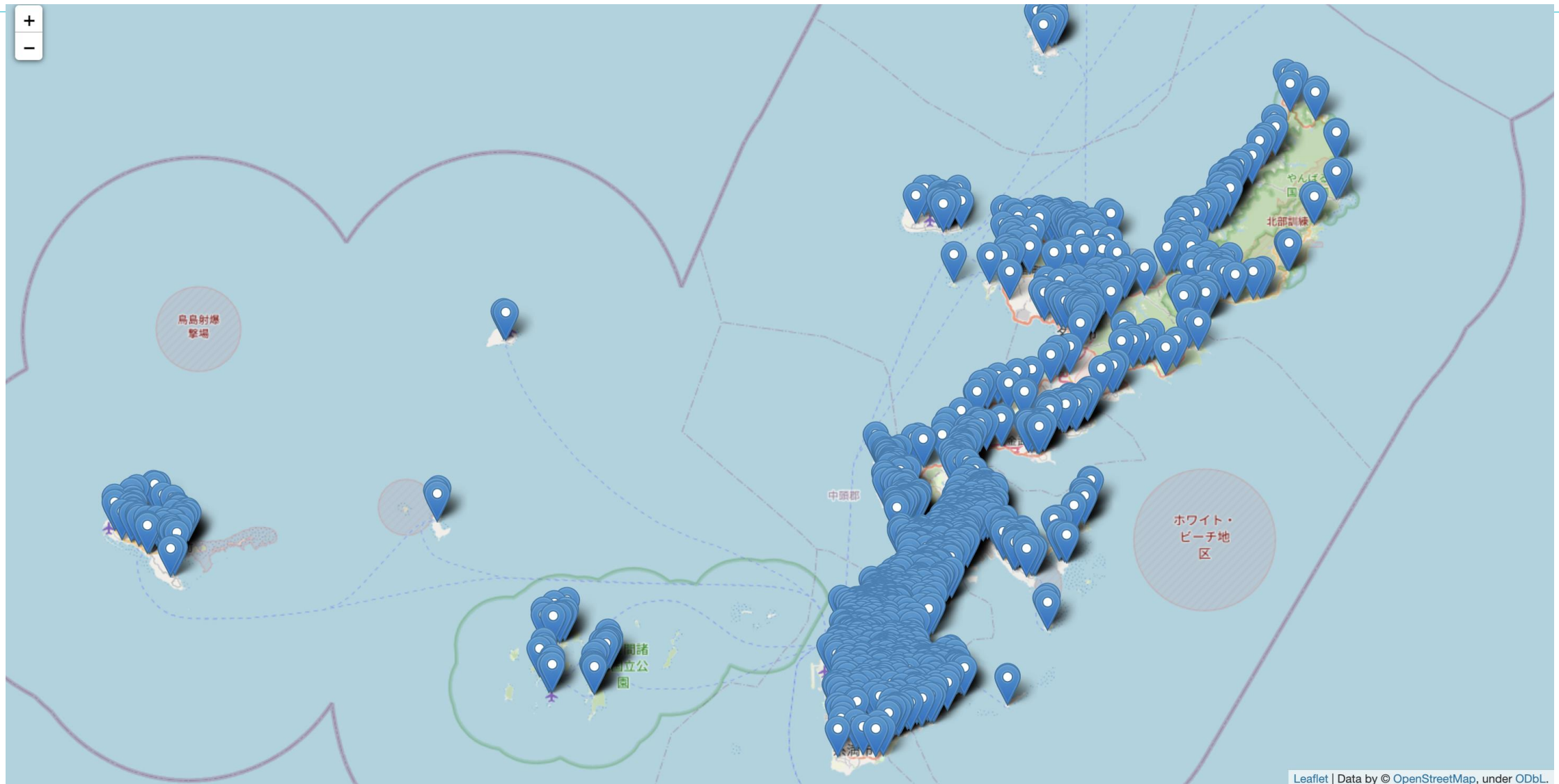


ChatGPT
旧Code Interpreter

命令：このgeojsonファイルを地図上にプロットしてHTML形式で保存してほしい

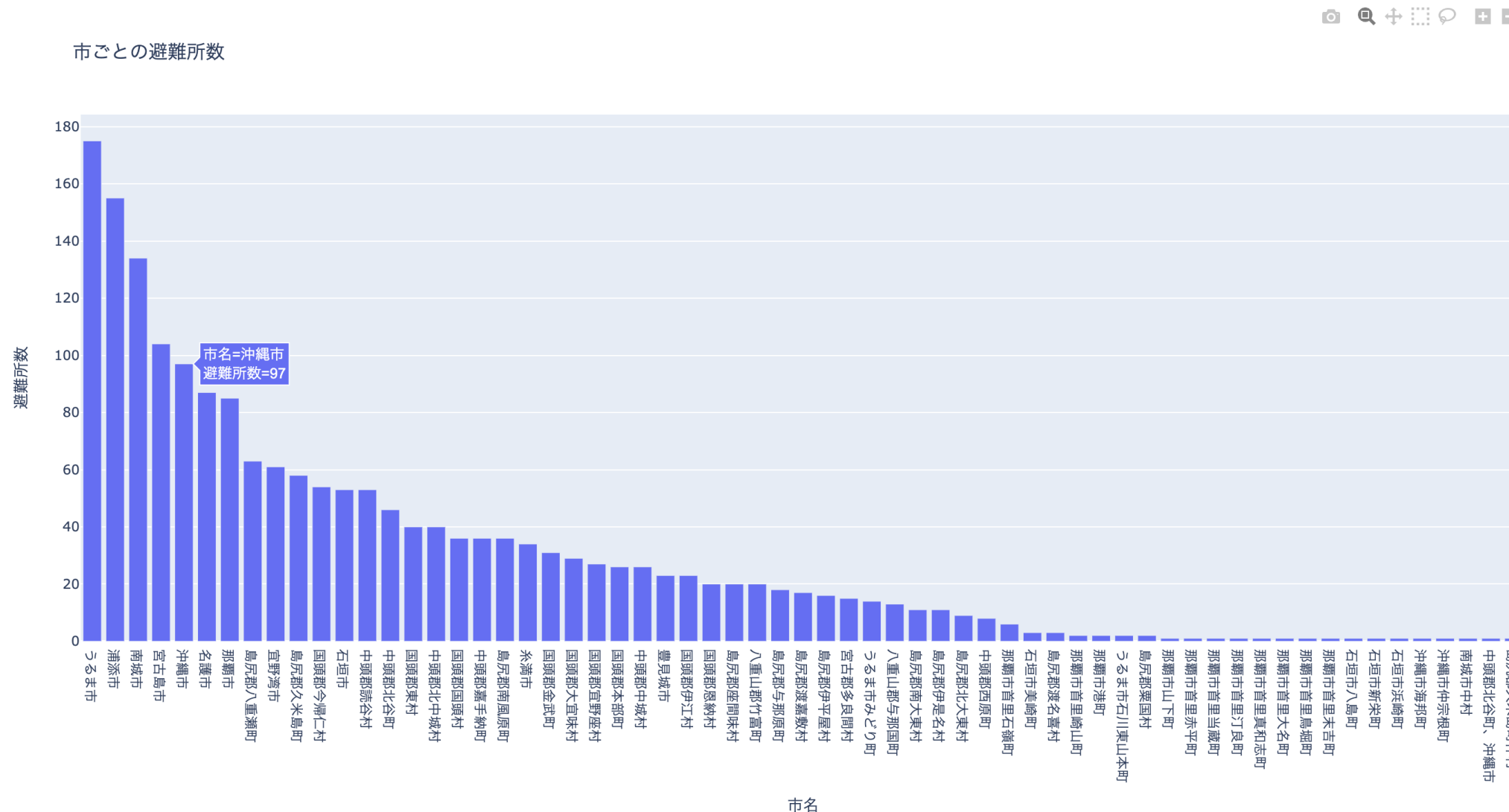
```
1 {  
2   "type": "FeatureCollection",  
3   "name": "KS-META-P20-12_47",  
4   "crs": { "type": "name", "properties": { "name": "urn:ogc:def:crs:EPSG::4612" } },  
5   "features": [  
6     { "type": "Feature", "properties": { "行政区域": "47201", "名称": "J A会館", "住所":  
       "沖縄県那覇市", "施設の種類": "緊急避難場所", "収容人数": -1, "施設規模": -1, "地震災害":  
       0, "津波災害": 1, "水害": 0, "火山災害": 0, "その他": 0, "指定なし": 0, "レベル": 3,  
       "備考": null, "緯度": 26.202377085199998, "経度": 127.685837541, "NO": "1" },  
       "geometry": { "type": "Point", "coordinates": [ 127.685837541,  
       26.202377085199998 ] } },  
7     { "type": "Feature", "properties": { "行政区域": "47201", "名称": "がじゃんびら公園",  
       "住所": "沖縄県那覇市金城1丁目", "施設の種類": "一時避難場所", "収容人数": -1, "施設規模":  
       17000, "地震災害": 0, "津波災害": 0, "水害": 0, "火山災害": 0, "その他": 0, "指定なし":  
       1, "レベル": 1, "備考": null, "緯度": 26.20348932, "経度": 127.66724343, "NO": "2"  
       }, "geometry": { "type": "Point", "coordinates": [ 127.66724343, 26.20348932 ]  
       } },  
   ]  
}
```

G空間センター国土数値情報（避難施設）-沖縄県
<https://www.geospatial.jp/ckan/dataset/ksj-p20-47>



OpenStreetMap上に避難所のピンが立ったHTMLが生成された
※実際はChatGPTといくつかのやり取りが発生

先ほどのデータを市ごとに避難所数を集計してグラフ化を命令



Plotlyを使ったインタラクティブなグラフのHTMLページが生成された
※実際はChatGPTといくつかのやり取りが発生

1. オープンデータにおける生成AIの活用可能性は大いにある
2. しかし、ハルシネーション(幻覚)が発生する可能性もあるため最終的には人間のチェックが必要となる
3. 100%の仕事ができなかったとしても20%でも30%でも人間の仕事を手伝ってくれると考えれば生成AIを活用している意味はある

ご清聴ありがとうございました

ITで、社会の願い叶えよう。

