# CHAMELEON

**FOR OUR SMARTER WORLD**

AI Dataset Audit – Investigation Findings

Hamish Burnett – (222282244)

# Contents

# Executive Summary

Large datasets are used to train AI models. These datasets can contain sensitive information, which may be leaked by the applications that use this data. Measures to prevent data leakage, include ensuring that training datasets do not contain sensitive information, sensitive information being removed from the dataset, and adding noise to the dataset, to prevent identification of individual records.

Following the audit of the datasets used by the EVCFLO Team, it was found that all datasets were suitable, and did not contain any sensitive data. Therefore, these datasets can be safely used, without any risk of leaking sensitive information.

# Audit Analysis

## Introduction

For an Artificial Intelligence (AI) algorithm to operate, a dataset is needed, to train the algorithm. The dataset can consist of millions of pieces of data, which are combined, to produce a meaningful output, and outcome. Care needs to be taken, to ensure that the algorithm ensures confidentiality, and does not reveal individual data points, from the data that it was trained on, when responding to queries (AI Exchange, n.d.-a). Specific data points which require attention to confidentiality, include information that could identify a person, or business secrets, which have been entered into the model (perhaps through a user entering a prompt, which contains a business secret) (AI Exchange, n.d.-d). There are several methods to ensure that confidentiality is maintained, including through removal of data that is not relevant to the algorithm (i.e. a person's name, phone number, or email address, may not be relevant, in an algorithm that analyses traffic patterns in a certain area) (AI Exchange, n.d.-c), obfuscation of data (i.e. add some noise to the data, which hides individual records, but still enables analysis of the whole dataset) (AI Exchange, n.d.-e), and only use data which was specifically collected for the purpose of developing an AI algorithm (i.e. if data was collected to verify the identity of individuals, this data (such as Driver's Licenses, and Passports), should not be used in an AI algorithm, to analyse vehicle traffic) (AI Exchange, n.d.-b). Consequences of leaking confidential data, may result in reputational damage, and legal action.

Thus, the datasets that are used within the Chameleon company, will be audited, to determine whether the AI models that are developed, are at risk of leaking sensitive information.

## Tools Used

GitHub was the sole tool that was utilised for this audit. GitHub contains the datasets which were used for developing the AI models.

## Scope of Audit

The scope of this audit, will be limited to the datasets used by the EVCFLO Team, which is developing EV Charger Forecasting & Location Optimisation algorithms. They use Machine Learning algorithms, to predict locations where EV chargers should be located (EVCFLO Team, 2024).

## Methodology

All of the datasets will be analysed, to identify whether any sensitive information is stored, and used within the AI algorithms. From an initial analysis, it was identified that the software that is developed, uses data to predict where charging locations should be located. It was also identified that the majority of data appears to be of locations where EV chargers are currently located. As this project is analysing charger locations, and plotting them on a map, it is believed that information about individual customers (names, phone numbers, email addresses, physical addresses, cars that they own, or the locations that particular individuals use, to charge their vehicle), should not be included in these datasets. Thus, particular attention will be given to personal data, and ensuring that the datasets do not contain any personal information. The datasets will be analysed, from the following link: https://github.com/Chameleon-company/EVCFLO/tree/main/datasets

## Results

### BEV.csv

The first dataset to be analysed, is the BEV.csv file, located in the *Australian Vehicle Specification Information* folder (SijuDENG, 2023a). A BEV, is a Battery Electric Vehicle, which runs on electricity (Lee, 2021). A screenshot of the data is shown below.

| | Model | Vehicle Type | Variant Details | Fast Charge Time (Minutes) | Range (km) | Energy Consumption (kWh/100km) | Listed Price ($AUD) |
|---|---|---|---|---|---|---|---|
| 1 | Model | Vehicle Type | Variant Details | Fast Charge Time (Minutes) | Range (km) | Energy Consumption (kWh/100km) | Listed Price ($AUD) |
| 2 | Audi e-tron | Large SUV | 2022 Audi e-tron 55 quattro | 85 mins (5%-80% charge, 50kW charger) | 459 | 23.2 | |
| 3 | Audi e-tron GT | Large Car | 2023 Audi e-tron GT Auto quattro MY23 | 23 mins (5%-80% charge, 270kW charger) | 540 | 19.2 | 180200 |
| 4 | BMW i4 | Medium Car | 2023 BMW i4 eDrive40 M Sport G26 Auto | 18 mins (10%-80% charge, 50kW charger) | 520 | 22.2 | 99900 |
| 5 | BMW i7 | Large Car | 2023 BMW i7 xDrive60 M Sport G70 Auto AWD | 21 mins (10%-80% charge, 50kW charger) | 625 | 22.2 | 306900 |

This data appears to contain the various different models of cars, that are Battery Electric Vehicles (BEV, or EV, for short). The information contained in this dataset, include the brand, and model (i.e. Audi e-tron), the vehicle type (i.e. Large SUV), the variant details (2022 Audi e-tron 55 quattro), the charging time (85 minutes), range (459 km), energy consumption (23 kWh/100 km), and the selling price (Not available for this model). None of the data appears to be sensitive personal information, or other information, that would be deemed at risk of confidentiality breaches. As this dataset relates to cars, that use EV charging stations, this data is relevant to the project, and should be kept. Thus, there are no issues with this data, and this data should be kept, without any modifications.

### PHEV.csv

The second dataset to be analysed, is the PHEV.csv file, located in the *Australian Vehicle Specification Information* folder (SijuDENG, 2023d). A PHEV, is a plug-in-hybrid electric vehicle. This means that it uses both traditional fuel (i.e. petrol, diesel), as well as electricity (in the form of a battery).

It contains the same information as the BEV dataset, except for PHEV vehicles, rather than BEVs. Thus, as the data is relevant to analysing EV charger locations, and is the same as the BEV data, this dataset should be kept, and does not pose any confidentiality risks.

### Cleaned_Australian_EV_Charging_Stations.csv

The next dataset to be analysed, is the *Cleaned_Australian_EV_Charging_Stations.csv* file, located in the Australian EV Charging Stations folder (See the following image) (jenniferyau, 2023a). This file contains the locations of EV chargers located around Australia, as well as accompanying information, for each charger. This CSV file contains a Location Name for each charger (Chargefox Charging Station), the precise location of the charger (Latitude and Longitude), and the address, town, city, and post code. Information specific to the individual charger is also included, which contains the type of plug that the charger caters for, the amount of chargers available at this location, and nearby amenities, such as hospitals, supermarkets, and parks.

All of this data is relevant to the goal of developing AI algorithms to predict locations of chargers, and none of the information has been found to contain sensitive information. As a result, all of the fields in this file, should be kept, and no modifications are required.

| | Location Name | Latitude | Longitude | Town | Postal Code | City | Address | Plugs_Type2 | Plugs_Three_Phase | Plugs_CHAdeMO | Plugs_CCS/SAE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Chargefox Charging Station | -35.72671957 | 145.6593541 | Cobram | 2714 | N.A | 154 Campbell Rd | 1 | | 0 | 0 |
| 1 | Secure Parking | -37.832108 | 144.969126 | N.A | 3205 | Melbourne | 15-29 Bank Street, South Melbourne VIC 3205, Australia | 0 | 0 | 0 | 0 |
| 2 | Lonsdale St | -37.81343708 | 144.955934 | N.A | 3000 | Melbourne | 535 Little Lonsdale St, Melbourne VIC 3004, Australia | 0 | 0 | 0 | 0 |
| 3 | Sorrento Community Centre | -38.33856 | 144.737043 | N.A | 3943 | Melbourne | 860 Melbourne Rd, Sorrento VIC 3943, Australia | 0 | 0 | 4 | 4 |
| 4 | CTR Whittlesea | -37.51059 | 145.118412 | N.A | 3757 | Melbourne | 74A Church St, Whittlesea VIC 3757, Australia | 0 | 0 | 2 | 2 |
| 5 | Rye & Sorrento Camping | -38.320637 | 144.718201 | N.A | 3943 | Melbourne | 1380 Point Nepean Rd Portsea VIC 3944, Australia | 0 | 0 | 0 | 0 |

| Plugs_Tesla | Plugs_J-1772 | Plugs_Caravan_Mains_Socket | Plugs_wall_AU/NZ | Power 1 | charging_stations | Nearby EVStations | Hospitals | Parks | Restaurants | Malls | Supermarkets |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 21 | | 21 | 15 | 1 | 18 | 20 | 4 | 20 |
| 2 | 0 | 0 | 0 | | 2 | 20 | 2 | 18 | 20 | 20 | 20 |
| 0 | 0 | 0 | 0 | 50.065 | 8 | 1 | 0 | 5 | 20 | 0 | 6 |
| 0 | 0 | 0 | 0 | 50.065 | 4 | 1 | 0 | 9 | 20 | 0 | 1 |

This file appears to be the same as another file (*New_Australian_EV_Charging_Stations_v2.csv*), which contains additional information, such as whether the charger is near coffee shops, or schools (jenniferyau, 2023b). The dataset that was analysed previously, also looks to be cleaned, in comparison to this dataset. There are no fields that are of concern, within this dataset, and this file is suitable to be kept.

## Station_Compiler_Location_T2_2023.csv

The next file to be analysed, is the *Station_Compiler_Location_T2_2023.csv* file, which is located in the *Station Compiler Location T2 2023* folder (See following image) (jamiecdavidson, 2023e). This file contains the locations of Service Stations, within Australia. It contains the exact address (i.e. 9 Murray Rose, Sydney), as well as the Latitude and Longitude of chargers. This file is relevant to analysing data for EV charging locations, and all of the data is suitable for this purpose. Therefore, this file should be kept, and no modifications are needed, to ensure confidentiality.

| | Service_Station_Location | Latitude | Longitude |
|---|---|---|---|
| 1 | Service_Station_Location | Latitude | Longitude |
| 2 | 9 Murray Rose,SYDNEY OLYMPIC PARK NSW 2127,AUSTRALIA | -33.84589758 | 151.0697681 |
| 3 | 76 Cowper St,WALLSEND NSW 2287,AUSTRALIA | -32.90277958 | 151.669841 |
| 4 | Hunter Valley Gardens 2090 Broke Road,POKOLBIN NSW 2320,AUSTRALIA | -32.77392928 | 151.293163 |
| 5 | Cnr Hume Highway & Bessemer Street,MITTAGONG NSW 2575,AUSTRALIA | -34.4495405 | 150.4429263 |
| 6 | 140 Queen St,BERRY NSW 2535,AUSTRALIA | -34.77593897 | 150.7005415 |

## Charger_Compiler_Type_T2_2023.csv

The next dataset to be analysed, is the *Charger_Compiler_Type_T2_2023.csv* file (jamiecdavidson, 2023b). This contains information about the precise location (latitude and longitude), of chargers, and the charger type (i.e. Type2, or Tesla charger). This information is relevant to plotting the locations of EV chargers, and all fields are suitable for this purpose. Therefore, this file can remain as it is, and no modifications are needed, to prevent leakage of sensitive data.

| 1 | Latitude | Longitude | charger_type |
|---|---|---|---|
| 2 | -35.72671957 | 145.6593541 | Type2 |
| 3 | -38.451339 | 145.240243 | Type2 |
| 4 | -38.129308 | 144.345207 | Type2 |
| 5 | -38.282256 | 145.125377 | Type2 |
| 6 | -16.8476624 | 145.6950452 | Type2 |

## Forecasting_global_charging_points_in_2024.csv

The next file to be audited, is the *Forecasting_global_charging_points_in_2024.csv* file (See following image) (SijuDENG, 2023b). This file contains the predictions of how many charging points will be in each suburb, for the year of 2024. The data included in this file, includes the number of predicted charging locations (i.e. 29), the country (Australia), and the suburb, or city (i.e. Altona). This information is essential for predicting where to install charging locations in the future, and as such, is highly relevant. The individual data points, are relevant, and none of the data points need to be modified, or removed. As such, this dataset is deemed to be acceptable, and does not need any modifications, as it already complies with confidentiality guidelines.

| 1 | Predicted Charging Points | Country | State/City |
|---|---|---|---|
| 2 | 29.355902004453927 | Australia | Altona |
| 3 | 7.338975501113482 | Australia | Altona North |
| 4 | 7.338975501113482 | Australia | Anglesea |
| 5 | 7.338975501113482 | Australia | Armidale |
| 6 | 14.677951002226964 | Australia | Ashwood |

## Global_EV_Data_26042023.csv

The next dataset to be analysed, is the *Global_EV_Data_26042023 data.csv* dataset (See the following image) (SijuDENG, 2023c). This file appears to show the amount of EV products that were purchased, within a country, for multiple countries. It shows the country (i.e. Australia), whether the data is historical data, or an estimate for the year (i.e. the *category* column), what type of sale was made (whether it was of EV sales, or sales on the stock market, charging points, and electricity demand), the type of vehicle that the purchase was relating to (i.e. cars, or buses, trucks), the drive system of the vehicle (whether it was an EV, or a PHEV), the year that the sales were completed, and the quantity (value variable) (i.e. how many a certain country purchased). These are all related to the topic of EVs, and being able to effectively plan for future demand, to make accurate models. All of the data is relevant, and none of the data needs to be modified or removed. As such, this file can be kept in its entirety, for the purposes of using it in AI models.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | region | category | parameter | mode | powertrain | year | unit | value |
| 2 | Australia | Historical | EV sales | Cars | BEV | 2011 | Vehicles | 49 |
| 3 | Australia | Historical | EV stock | Cars | BEV | 2011 | Vehicles | 49 |
| 4 | Australia | Historical | EV sales share | Cars | EV | 2011 | percent | 0.0065 |
| 5 | Australia | Historical | EV stock share | Cars | EV | 2011 | percent | 0.00046 |
| 6 | Australia | Historical | EV stock share | Cars | EV | 2012 | percent | 0.0028 |
| 7 | Australia | Historical | EV sales share | Cars | EV | 2012 | percent | 0.029999999329447746 |

## Elektrick%C3%A9_nab%C3%ADjec%C3%AD_stanice_pro_auta___EV_charging_points.csv

The next file to be analysed, is the
Elektrick%C3%A9_nab%C3%ADjec%C3%AD_stanice_pro_auta___EV_charging_points.csv file, which
contains information about charging locations in the Czech Republic (jamiecdavidson, 2023c). This
file contains information about the operator information of individual chargers, and the associated
website, whether the public can freely use the charger, or if they are reserverved for certain
individuals (i.e. staff), whether membership or a key is needed, the payment information, exact
addresses, including latitude and longitude, and the connector information.

Although this dataset contains information such as whether anyone can use it, or if it is only used by
employees, membership requirements, and key requirements, this data does not identify an
individual that has all of these components. This may make it possible to identify a person,
particularly if it is known that only a certain person is allowed to use that charger (i.e. an employee,
with a membership, at a certain location). However, this would not be able to be used to identify a
person, as sensitive information about individual's is not included in this dataset. As this information
is directly related to EVs, and no information was found to be sensitive, this file is given the all-clear,
to be used, for analysis.

| | X | Y | ObjectId | OperatorInfo | OperatorInfoWebsite | UsageType | MembershipRequired | PayAtLocation | KeyRequired | StatusType |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | |
| 2 | 1838534.8988 | 6316293.0025 | 1 | Tesla (Tesla-only charging) | http://www.teslamotors.com | Public - Membership Required | True | False | True | Operational |
| 3 | 1840113.7264 | 6322320.5019 | 2 | E.ON (CZ) | http://www.ekobonus.cz/ekologicka-doprava/elektromobilita | Public | Unknown | Unknown | Unknown | Operational |
| 4 | 1851397.7435 | 6314134.9575 | 3 | (Unknown Operator) | Unknown | Unknown | Unknown | Unknown | Unknown | Unknown |
| 5 | 1846761.2948 | 6314373.9084 | 4 | e-Mobilita Brno (cz) | https://www.emobilitabrno.cz/en/chargers | Public - Notice Required | False | False | False | Operational |
| 6 | 1845580.1818 | 6313454.7096 | 5 | ČEZ | http://www.elektromobilita.cz/ | Public - Membership Required | True | False | True | Operational |
| 7 | 1848494.7521 | 6308103.3452 | 6 | E.ON (CZ) | http://www.ekobonus.cz/ekologicka-doprava/elektromobilita | Public | Unknown | Unknown | Unknown | Operational |

| AddressSpecs | AddressInfo | City | latitude | Longitude | Connections |
|---|---|---|---|---|---|
| Maximus Resort | Hrazni 327/4a | Brno | 49.243578 | 16.51584 | [{'Amps': None, 'Voltage': None, 'PowerKW': 17.0, 'Quantity': 1, 'ConnectionType': 'Tesla (Model S/X)', 'StatusType': 'O |
| Golfový resort Kaskáda | Na Golfu 1772 | Kuřim | 49.2789142416232 | 16.5300228490043 | [{'Amps': 32, 'Voltage': 400, 'PowerKW': 22.0, 'Quantity': 1, 'ConnectionType': 'Type 2 (Socket Only)', 'StatusType': 'Op |
| Jiří Hořice | Seifertova 519/50 | Brno | 49.2309203 | 16.6313889 | [{'Amps': 16, 'Voltage': 400, 'PowerKW': 11.0, 'Quantity': '1', 'ConnectionType': 'CEE 5 Pin', 'StatusType': 'Unknown', 'C |
| Knihovna Jiřího Mahena | Palackého třída 164 | Brno | 49.2323219916187 | 16.5897389727613 | [{'Amps': 40, 'Voltage': 400, 'PowerKW': 27.7, 'Quantity': 2, 'ConnectionType': 'Type 2 (Socket Only)', 'StatusType': 'Op |
| 3079/99f Purkyňova | 3079/99f Purkyňova | Brno | 49.22692973234 | 16.5791288540499 | [{'Amps': None, 'Voltage': None, 'PowerKW': 50.0, 'Quantity': 1, 'ConnectionType': 'CHAdeMO', 'StatusType': 'Operati |
| Best Western Premier Hotel International Brno, Hotel Garage, Besedni ulice | Besední | Brno | 49.1955255489139 | 16.6053108845797 | [{'Amps': 32, 'Voltage': 400, 'PowerKW': 22.0, 'Quantity': 2, 'ConnectionType': 'CEE 5 Pin', 'StatusType': 'Operational', ' |

## Melbourne_Landmarks_and_POIs.csv

The next dataset to be analysed, is the *Melbourne_Landmarks_and_POIs.csv* file, which contains information about landmarks, and locations of interest, in Melbourne (west557, 2022a). This includes a theme of a location (i.e. Leisure, Place of Worship), a sub theme (i.e. Church, Garden), the name of the attraction (i.e. Flemington Bridge Railway Station), and the precise location of the attraction, in the form of latitude and longitude. This data is potentially relevant to the task of analysing EV charger locations, as users may wish to search for chargers near particular facilities, such as churches, or parklands. All of the data is relevant, and none of the data is deemed to be confidential in nature. Therefore, this dataset is allowed to be used, and does not require any modifications, to ensure confidentiality.

| | Theme | Sub Theme | Name | Latitude | Longitude |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | Transport | Railway Station | Flemington Bridge Railway Station | -37.78816459 | 144.9392778 |
| 3 | Mixed Use | Retail/Office/Carpark | Council House 2 (CH2) | -37.81425914 | 144.9666384 |
| 4 | Place Of Assembly | Library | The Melbourne Athenaeum Library | -37.81488558 | 144.9672913 |
| 5 | Leisure/Recreation | Informal Outdoor Facility (Park/Garden/Reserve) | Carlton Gardens South | -37.80606846 | 144.9712665 |
| 6 | Place of Worship | Church | St Francis Church | -37.81188478 | 144.9624226 |

## SA Data – registered-vechicles-by-postcode-at-30-june-2022.csv

The next dataset to be analysed, is the *SA Data – registered-vehicles-by-postcode-at-30-june-2022.csv* dataset (west557, 2022b). This file contains information about the details of cars that were registered in South Australia, and includes the owners post code, the Vehicle Make (i.e. BMW, Holden), the Vehicle Body Type (i.e. Sedan, Caravan, Trailer), and the Count of Vehicle ROIDs. This ROIDs is unfamiliar, and it was unable to be identified. The source of the data (https://data.sa.gov.au/data/dataset/registered-vehicles-by-postcode ) was also analysed, which seemed to indicate the amount of cars, of this car type, and car body, in a particular suburb. This dataset does not identify individuals, the cars that an individual owns, and does not contain any other information that would be classified as sensitive. This data also meets the Data Obfuscation strategy, listed in the introduction, which is to show what the community owns, but not be able to identify an individual, through grouping cars by their postcode, and not identifying an individual, with their car.

| | Owner PostCode | Vehicle Make | Vehicle Body Type | Count of Vehicle ROIDs |
|---|---|---|---|---|
| 7 | | | | |
| 8 | 800 | BMW | SEDAN | 1 |
| 9 | 800 | CABANA | CARAVAN | 1 |
| 10 | 800 | COROMAL | CARAVAN | 1 |
| 11 | 800 | HOLDEN | STATION WAGON | 1 |
| 12 | 800 | HOMEMADE | TRAILER | 1 |
| 13 | 800 | MITSUBISHI | STATION WAGON | 1 |

## alt_fuel_stations (Apr 27 2023).csv

The next dataset to be analysed, is the *alt_fuel_stations (Apr 27 2023).csv* file, which contains a large amount of information, for fuel stations in the US (jamiecdavidson, 2023a). This is quite a large dataset, and contain many different fields. As it is so large, it will be analysed in several sections.

The dataset contains a fuel type (ELEC), A Service Station Name, and accompanying information, on locations, and directions to get to the location. The phone number for each station is also listed,

along with a status code, availability times of the charger, and whether the charger is available for the public to use. It also stores the cards that can be used to pay for it. All of this data, is necessary for the development of algorithms to locate EV charging locations, and none of the data contains any sensitive information. As such, this data can be used, and does not need to be modified.

The data also includes specific information about the fuel that it offers, including BD Blends, the EV charger that it is, whether it is a fast charger, and whether it is part of an EV network, and the cost of EV charging, at individual stations. Information that relates to the last time that the charger was confirmed, and the last updated time, is also included. The data includes details about the federal agency that oversees the charger, and provides details about whether hydrogen charging, LPG, E85 Pumps, and EV Connector types, are supported, or not. The type of the facility, where the charger is located (i.e. Hospital, Grocery Store) is also included. All of this data is appropriate for use, within this algorithm, and does not need to be modified, for confidentiality reasons.

Some of the data has been duplicated, and is in French as well.

It has been found, that all data in this dataset is necessary, and appropriate, and that none of the data contains information that would have confidentiality issues. Therefore, this dataset can be used, and does not need to be adapted, as it already ensures confidentiality.

| Fuel Type ( | Station Na | Street Add | Intersectic | City | State | ZIP | Plus4 | Station Ph | Status Coc | Expected L | Groups Wi | Access Da | Cards Acc | BD Blends | NG Fill Typ | NG PSI | EV Level1 I | EV Level2 I | EV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ELEC | Los Angele | 1201 S Fig | West hall a | Los Angele | CA | 90015 | | 213-741-1 | E | | Public | 5:30am-9pm; pay lot | | | | | | 7 | |
| ELEC | California | 9530 Telstar Ave | | El Monte | CA | 91731 | | 626-575-6 | E | | Public | 24 hours daily | | | | | | 3 | |
| ELEC | Scripps Gr | 10666 N Tc | Patient Pa | La Jolla | CA | 92037 | | | E | | Public | 24 hours daily | | | | | | 1 | |
| ELEC | Galpin Mot | 15421 Roscoe Blvd | | Sepulveda | CA | 91343 | | 800-256-6 | E | | Public - Ca | Dealership business hours | | | | | | 2 | |
| ELEC | Galleria at | 1299 Galleria at Tyler | | Riverside | CA | 92503 | | 951-351-3 | E | | Public | 6am-12am daily | | | | | | 4 | |
| ELEC | City of Pas | 150 E Holly St | | Pasadena | CA | 91103 | | 626-744-7 | E | | Public | 24 hours daily; pay lot | | | | | | 16 | |

### ladestationen_elektrofahrzeuge.csv

The final dataset to be analysed, is the ladestationen_elektrofahrzeuge.csv dataset, which contains data about EV chargers in Germany (jamiecdavidson, 2023d). Even though the dataset is in German, it is important to audit this dataset, to ensure that confidentiality of the data is maintained. With the dataset in German, this will present some challenges, but Google Translate will be used, for translations. The dataset contains the precise location of charging stations (in the form of latitude and longitude), an ID number (uuid), a city name (gemeindeverband_name) (i.e. Rostock, Hanse- und Universitätsstadt, which translates to Rostock, Hanseatic and university city, which is a city in Germany), and ID, community name (gemeinde_name – i.e. Rostock, Hanseatic and university city; which translates to a city in Germany), community ID number (gemeindeverband_schluessel), and a community part name, and key (gemeindeteil_name, gemeindeteil_schluessel). The dataset also contains data which relates to the streetname (strasse_schluessel), street number (hausnummer), post code (postleitzahl), and organisation (bezeichnung) which manages the EV Charger. The dataset also includes the operating mode of the EV charger (betriebsart), and number of charging points (anzahl_ladepunkte) at a particular location.

It was found that all of the fields are relevant to EV Chargers, and that none of the fields raised any confidentiality issues. Therefore, this dataset can continue to be used, in its current state, and does not need any modifications.

## Conclusion

After performing a strong analysis, on the datasets that the EVCFLO Team uses, it was found that there were no issues regarding confidentiality, of training data. All the data has related directly to EV Charging locations, and none of the data contained personal information, such as individual names, contact information, or any information about the cars that an individual owns. Therefore, these datasets can be safely used, for data analysis, without risk of leaking confidential information.

## References

AI Exchange (n.d.-a) *AI Security Overview*: OWASP, accessed 2024.
https://owaspai.org/docs/ai_security_overview/

AI Exchange (n.d.-b) *#ALLOWEDDATA*: OWASP, accessed 2024.
https://owaspai.org/docs/1_general_controls/#alloweddata

AI Exchange (n.d.-c) *#DATAMINIMIZE*: OWASP, accessed 2024.
https://owaspai.org/docs/1_general_controls/#dataminimize

AI Exchange (n.d.-d) *#DEVDATAPROTECT*: OWASP, accessed 2024.
https://owaspai.org/docs/3_development_time_threats/#devdataprotect

AI Exchange (n.d.-e) *#OBFUSCATETRAININGDATA*: OWASP, accessed 2024.
https://owaspai.org/docs/1_general_controls/#obfuscatetrainingdata

EVCFLO Team (2024) *EVCFLO*: GitHub, accessed 2024.: https://github.com/Chameleon-company/EVCFLO/tree/main

jamiecdavidson (2023a) *alt_fuel_stations (Apr 27 2023).csv*: GitHub, accessed 2024.
https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T1_2023/US_Alternative_Fueling_Stations/alt_fuel_stations%20(Apr%2027%202023).csv

jamiecdavidson (2023b) *Charger_Compiler_Type_T2_2023.csv*: GitHub, accessed 2024.
https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Charger_Compiler_Type_T2_2023.csv

jamiecdavidson (2023c) *Elektrick%C3%A9_nab%C3%ADjec%C3%AD_stanice_pro_auta___EV_charging_points.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T1_2023/Czech_Republic/Elektrick%25C3%25A9_nab%25C3%25ADjec%25C3%25AD_stanice_pro_auta___EV_charging_points.csv

jamiecdavidson (2023d) *ladestationen_elektrofahrzeuge.csv*: GitHub, accessed 2024.
https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T1_2023/Germany/ladestationen_elektrofahrzeuge.csv

jamiecdavidson (2023e) *Station_Compiler_Location_T2_2023.csv*: GitHub, accessed 2024.
https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Station_Compiler_Location_T2_2023/Station_Compiler_Location_T2_2023.csv

jenniferyau (2023a) *Cleaned_Australian_EV_Charging_Stations.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Australian_EV_Charging_Stations_T2_2023/Cleaned_Australian_EV_Charging_Stations.csv

jenniferyau (2023b) *New_Australian_EV_Charging_Stations_v2.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Australian_EV_Charging_Stations_T2_2023/New_Australian_EV_Charging_Stations_v2.csv

Lee, C. (2021) *What is EV, BEV, HEV, PHEV? Here's your guide to types of electric cars*: MarketWatch, accessed 2024. https://www.marketwatch.com/story/what-is-ev-bev-hev-phev-heres-your-guide-to-types-of-electric-cars-11617986782

SijuDENG (2023a) *BEV.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Australia_Vehicle_specification_information_13042023/BEV.csv

SijuDENG (2023b) *Forecasting_global_charging_points_in_2024.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Forecasting_global_charging_points_in_2024.csv

SijuDENG (2023c) *Global_EV_Data_26042023 data.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Global_EV_Data_26042023%20data.csv

SijuDENG (2023d) *PHEV.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T2_2023/Australia_Vehicle_specification_information_13042023/PHEV.csv

west557 (2022a) *Melbourne_Landmarks_and_POIs.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T3_2022/Melbourne_Landmarks_and_POIs.csv

west557 (2022b) *SA Data - registered-vehicles-by-postcode-at-30-june-2022.csv*: GitHub, accessed 2024. https://github.com/Chameleon-company/EVCFLO/blob/main/datasets/T3_2022/SA%20Data%20-%20registered-vehicles-by-postcode-at-30-june-2022.csv