

Using LLMs to 'Break Down' the Black Box

David Schmitt, Executive Vice President
2025 Modeling Mobility Conference
September 15th, 2025



Topics

What are Large Language Models (LLMs)?

What are GPTs?

How can both be used to help break down the travel model black box?

Examples of travel model-related GPTs

LLMs & GPTs...What are they?

- Large Language Models (LLMs)
 - An AI system trained on massive amounts of data so it can understand and generate human-like language
 - LLMs learn patterns, facts, and relationships in language, enabling them to answer questions, write text, summarize information.
 - Examples: ChatGPT, Claude, Grok, and many others
 - LLMs can also be the “engines” that power other many AI applications
- Generative Pre-Trained Transformers (GPTs)
 - A type of LLM that can be customized for specific purposes
 - Example customizations include:
 - By topic: law, medicine, etc.
 - By task: image generator, file converter, creative writing, transcribe audio, etc.
 - By style: formal, casual, etc.

GPTs: Application to Travel Modeling

How they work:

- 1) Identify **LLM** to use
- 2) Develop **Knowledge Base**
- 3) **Instruct** the GPT how to best address user queries
- 4) **Test** the GPT for reasonableness and reliability
- 5) **Distribute** GPT weblink to users

“Break Down” the Black Box

“Bridge the gap” between technical analyses & questions from non-technical stakeholders:

Input: ***natural language*** queries

Data: technical modeling data

Output: ***natural language*** answers

Examples:

1. Documentation reference
2. Model data queries
3. Data manipulator

GPT #1



Ask Me About CFRPM 7 GPT

By David Schmitt [in](#)

Provides answers to questions using the Central Florida Regional Planning Model Version 7 (CFRPM 7) model documentation. It does not provide detailed information about model results at this time. Please contact info@insightcinc.com with any questions.

Which counties are in CFRPM 7?

What software does CFRPM 7 use?

How was the employment data verified?

How were the socio-demographic...

Documentation reference for:
Central Florida Regional Planning Model 7
(CFRPM 7)

Ask Me About CFRPM 7

Knowledge Base & Instructions

Knowledge Base

- User's guide
- Model description report
- Model validation report
- Data dictionary (selected file formats)
- Training materials
- Facts list

Knowledge base materials converted to Markdown format

Instructions

- Response scope
- Response style & quality
- Security
- Follow-up questions

Instructions provided in Word

ChatGPT - Ask Me About CFRP! x +

chatgpt.com/g/g-q7pvbsw0q-ask-me-about-cfrpm-7-gpt?model=gpt-5

Ask Me About CFRPM 7 GPT 5 v

Ask Me About CFRPM 7 GPT

By David Schmitt

Provides answers to questions using the Central Florida Regional Planning Model Version 7 (CFRPM 7) model documentation. It does not provide detailed information about model results at this time. Please contact info@insightcinc.com with any questions.

Which counties are in CFRPM 7?

What software does CFRPM 7 use?

How was the employment data verified?

How were the socio-demographic variables...

Ask anything

7

GPT #2



Ask Me About CFRPM 7 Data GPT

By David Schmitt 

Provides answers to questions about the Central Florida Regional Planning Model Version 7 traffic results for the years 2015-2045 (in 5 year increments). Currently, traffic data can be provided for intersections or roadway segments.

What roadway segments have the top 5 larges...

What is the volume at the intersection of ...

Provides access to roadway segment & intersection volumes for 2015-2045 by time period

Offers ability to visualize traffic data with download

Ask Me About CFRPM 7 Data

Knowledge Base & Instructions

Knowledge Base

- Intersection data (CSV)
 - Roadway names, approach direction, area & facility types, county, lanes, speeds, AM / MD / PM / NT / daily volumes for 2015-2045
- Roadway segment data (CSV)
 - Same data organized by roadway segments

Instructions: general data analysis

Roadway1	SR 44
Roadway2	CR 143
Direction_in_2015	EB
NumberOfLanes_in_2015	2
PostedSpeedInMilesPerHour_in_2015	55
AreaType_in_2015	31: Residential Area of Urbanized Areas
FacilityType_in_2015	23: Divided Arterial Class I
County_in_2015	SUMTER
MPO_in_2015	Lake Sumter
DistanceInMiles_in_2015	0.19
DailyTrafficVolume_AMPeak_in_2015	1,432
DailyTrafficVolume_Midday_in_2015	3,435
DailyTrafficVolume_PMPeak_in_2015	3,095
DailyTrafficVolume_Evening_in_2015	2,757
DailyTrafficVolume_Daily_in_2015	10,719
DailyCapacity_AMPeak_in_2015	3,700
DailyCapacity_Midday_in_2015	6,300
DailyCapacity_PMPeak_in_2015	3,800
DailyCapacity_Evening_in_2015	5,800

ChatGPT - Ask Me About CFRP! x +

chatgpt.com/g-g-PIAh6nTAV-ask-me-about-cfrpm-7-data-gpt?model=gpt-5

Ask Me About CFRPM 7 Data GPT 5 v

Ask Me About CFRPM 7 Data GPT

By David Schmitt

Provides answers to questions about the Central Florida Regional Planning Model Version 7 traffic results for the years 2015-2045 (in 5 year increments). Currently, traffic data can be provided for intersections or roadway segments.

What roadway segments have the top 5 largest 2025...

What is the volume at the intersection of SR 434 and...

Ask anything

GPT #3



GTFS GPT

By David Schmitt [in](#)

Provides GTFS basic editing with natural language query.
Upload a compressed (zipped) file of GTFS files to request information or modify the GTFS files.

General Transit Feed Specification:
standardized data format used by most transit agencies to describe their service (key input to FTA's STOPS model)

Provides GTFS basic editing with natural language query

Upload a compressed (zipped) file of GTFS files to request information or modify the GTFS files

GTFS GPT

Knowledge Base & Instructions

Knowledge Base

- GTFS.org reference & best practices
- Instructions for minor tasks:
 - Modifying headway
 - Modifying route type (see right: sample)
 - Removing agency
 - Add agency
 - Add stop/station
- Python scripts to handle advanced tasks:
 - Adding and extracting route(s)
 - Filtering sets of routes
 - Deleting routes
- Instructions: step-by-step (see right)

```
### PRIMARY ACTION: MODIFYING ROUTE_TYPE
**Objective:**  
Update the `route_type` value for a specific route within a GTFS feed by modifying the appropriate CSV files, primarily `routes.txt`, while ensuring data consistency and validation throughout the process.

### Step-by-Step Instructions for the Custom GPT
##### 1. **Initialization and Input Gathering**  
1. **Welcome and Context Setup:**  
2. **Request Required Files:**  
3. **Collect User Input Details:**  
##### 2. **Processing the GTFS Feed**  
1. **Parse the Feed Files:**  
2. **Locate the Target Route Record:**  
3. **Modify the `route_type`:**  
4. **Save the Modified File:**  
##### 3. **Consistency and Optional Additional Updates**  
1. **Review Dependent Files:**  
2. **User Query:**  
##### 4. **Validation and Packaging**  
1. **GTFS Feed Validation:**  
2. **Repackage the Feed (if applicable):**  
3. **Log and Report the Changes:**  
##### 5. **Final Output**  
1. Output the Updated Files:  
2. Summary Message
```

ChatGPT - GTFS GPT

chatgpt.com/g/g-67a1694f44648191a0777c9c728008c8-gtfs-gpt

GTFS GPT 4.0

GTFS GPT

By David Schmitt

Provides GTFS basic editing with natural language query. Upload a compressed (zipped) file of GTFS files to request information or modify the GTFS files.

Ask anything

13

Lessons Learned

Constructing a Travel Modeling-Oriented GPT

- Recommend re-casting model-related data into formats & structures better compatible for your GPT objective
 - Markdown, not Word or PDF
 - CSV or JSON, not Excel
 - Data designed for filtering & analysis, not for model processing
- Performance varies depending on outside factors
- Be honest if hallucinations (“drifts”) occur
- Instructions are crucial: be specific!
- Periodically update instructions & add functionality

Please provide your feedback to David.Schmitt@Insighttcinc.com

Ask Me about
CFRPM 7

Model documentation reference



Ask Me About
CFRPM 7 Data

Model data querying



GTFS
GPT

Transit network manipulation



Thank you!

