

# Can Artificial Intelligence (AI) Models use Public Data, Running Locally, Answer Real Estate Sales Related Queries?

Umesh J Jadunandan

*School of Business and Management, Texila American University, Guyana, South America*

## Abstract

*Artificial Intelligence (AI) is evolving rapidly; however, the Real Estate Industry is slow to adopt AI Technologies. From the Real Estate point of view, Artificial Intelligence may be defined as utilization of computing methodologies upon a large dataset of real estate related data, providing insight into prices or trends beyond human cognitive capabilities. In this publication, quires, using natural language, on a dataset using an appropriate AI model running on a local computer is discussed. The dataset selected for this analysis is obtained from the public records of the Palm Beach County Property Appraiser's website. Palm Beach County is in the State of Florida, USA. Here, all real estate transactions are recorded with the Public Property Appraiser. This data is available to the public, thus providing a reliable low, or no cost dataset.*

**Keywords:** *AI Model, AI Dataset Analysis, Data Acquisition, Property Data, Real Estate Prices Trends.*

## Introduction

Investing in real estate can be a lucrative opportunity, but it also comes with risks and challenges. As AI technology continues to evolve, its role in real estate acquisition and sale will expand further. This publication examines the availability of AI models that can utilize Publicly available real estate sales data and ability to run the AI model on a person computer. AI-powered property models, with the aid of chatbots, will increasingly handle inquiries, property searches, all with a personalized buyers' preferences [21].

In all real estate transitions, the market dynamics, such as supply and demand, plays an important part in real estate valuation. Real Estate prices also vary with market conditions, fluctuate based on buyer/seller activity, economic conditions, and interest rates. Real Estate prices are also influenced by location, size, condition, and other market trends. Participating in real estate is not without its challenges. Some of the challenges include

market volatility, economic downturns, and changes in regulations can pose risks [4].

Palm Beach County, Florida, USA is selected to obtain the real estate sales data, since this data is readily downloadable from the County's website [18]. This is the largest County in the State of Florida. This County consists of over half million unique real estate parcels. The sale type selected in this study is limited to all residential sales above \$1000.00 USD over the last 5 years. The important feature of the dataset is that each parcel is assigned a property or parcel control number (PCN). This unique number does not change and is the key identifier that will allow the model to track the historical sales of each parcel over time. A successful AI Model that can utilize this data will mitigate risks and challenges with acquisition and sale of Real Properties [9].

## Download Sample

The data sample illustrated the PCN identifier, last sale price, the sale date, location and property size. However, some sales were recorded as \$10.00 whilst others were zero dollars. These data need to be

removed to ensure that the data represent the true market value of the real property sale. Any sale price below \$1000.00 were removed. Additional field that was available included the property addresses,

property use types, owner's names and property physical features. From the Property addresses, street names, municipality, zip codes can be quired as needed.

**Table 1.** Example of Cleaned Dataset, Palm Beach County Property Appraiser [1]

Parcel Number	Sale Price	Sale Date	Location	Municipality	Sq. Ft	Owner
38-43-44-21-15-254-0060	59000.00	8/19/2013	822 N D ST	L WORTH BEACH	1360	N
30-42-40-36-09-001-0140	242500.00	11/10/2004	105 W RIVERSIDE DR	JUPITER	1382	N
74-43-43-16-08-000-0810	260000.00	7/28/2023	1009 18TH ST	ISLAND PB	1024	N
08-43-45-09-12-011-0010	110000.00	8/18/2020	2960 N SEACREST BLVD	BOYNTON BEACH	660	Y
73-41-44-10-30-000-0480	260000.00	5/15/2018	1312 WATERWAY COVE DR	WELLINGTON	1352	Y
30-43-41-06-02-000-0321	112700.00	3/20/2012	104 YACHT CLUB DR	JUPITER	1331	N
50-43-43-35-00-002-0150	1411600.00	7/1/2017	1047 S OCEAN BLVD	ISLAND PB	6736	N
50-43-43-35-00-002-0150	1200000.00	5/21/2019	1047 S OCEAN BLVD	ISLAND PB	6736	N
52-42-42-16-14-000-0410	400000.00	8/12/2019	238 EAGLETON LAKE BLVD	PB GARDENS	2960	N
50-43-43-14-08-000-0053	4607500.00	12/11/2008	110 WELLS RD	ISLAND PB	4420	Y
00-42-45-32-03-000-1270	736000.00	10/31/2017	11110 STONEWOOD TRL	PB County	5595	Y
38-43-44-21-15-334-0030	95000.00	9/1/2015	1112 N D ST	L WORTH BEACH	920	N
00-42-43-24-00-000-7270	750000.00	1/9/2023	4664 SCHALL WAY	PB County	1847	N
40-43-45-03-19-006-0050	70800.00	6/18/2012	1116 PINE TREE DR	LANTANA	1523	Y
00-43-45-09-09-000-2640	322500.00	11/30/2022	1125 PEAK RD	PB County	1068	N
52-42-42-16-10-000-0130	445000.00	5/31/2013	113 COVENTRY PL	PB GARDENS	2269	Y
12-43-46-21-00-004-0060	1800000.00	11/3/2021	1171 S OCEAN BLVD	DELRAY BEACH	9761	N
40-43-45-04-18-008-0220	135000.00	1/15/2019	1171 PALAMA WAY	LANTANA	1998	Y

## Key Consideration

It is important to be careful to choose the model that is best suited for the specific needs. An important feature of the dataset is that each parcel is assigned a property or parcel control number. This unique number does not change and is the key identifier that will allow the model to track the historical sales or improvement of each parcel over time.

### Key Consideration:

1. Data type: Time series data [22].
2. Quires: Natural language.
3. Model complexity: Simple for a user without programming experience.

4. Interpretability: A model that's easy to explain and understand.

5. Computational resources: Home PC running Windows 11 Pro, i7-14700F Intel Processor @ 210 GHz and 32GB memory.

Data shown above (Table 1), is a times series data, since it is sales prices change over time [17]. The objective is that for the user to query the dataset using natural language. Natural language Intelligence processing (NLP) is where computer science and artificial intelligence bridges the gap between human and computer communication by enable natural language conversation [21]. These interfaces allow humans to interact with AI systems through spoken

or written languages [2]. There is no need-to-know complex programming languages or commands or queries. The human user can engage in conversation as if conversing with another human user. In addition, Large Language Model (LLM) [9], a type of AI systems can generate and process natural language using any amount of data. By asking, in natural language, tasks such as, property listing text summarization, question and answering on the dataset, perform analysis with novel generation and coherent responses.

Understanding your dataset and asking with an appropriate prompt can deliver profoundly amazing analysis for your dataset without complex formulas and programming. Prompt engineering is a technique used in natural language processing (NLP) to improve the performance of the language model by providing them with more context and information about the task in hand. It involves creating prompts, which are short pieces of text that provide additional information or guidance to the model, such as the topic or genre of the text it will generate. By using prompts, the model can better understand what kind of output is expected and produce more accurate and relevant results. In Llama 3 the size of the context, in terms of number of tokens is 128k.

### **Crafting Effective Prompts**

Crafting effective prompts is an important part of prompt engineering [25]. Here are some tips for creating prompts that will help improve the performance of your language model:

1. Be clear and concise: Your prompt should be easy to understand and provide enough information for the model to generate relevant output. Avoid using jargon or technical terms that may confuse the model.
2. Use specific examples: Providing specific examples in your prompt can help the model better understand what kind of output is expected. For example, if you want the model to generate a story about a particular topic, include a few sentences about the setting, characters, and plot.
3. Vary prompts: Using different prompts can help the model learn more about the task at hand and produce more diverse and creative output. Try using different styles, tones, and formats to see how the model responds.
4. Test and refine: Once you have created a set of prompts, test them out on the model to see how it performs. If the results are not as expected, try refining the prompts by adding more detail or adjusting the tone and style.
5. Use feedback: Finally, use feedback from users or other sources to continually improve your prompts. This can help you identify areas where the model needs more guidance and adjust accordingly.

### **Methodology**

In this study, the Llama model 3.1 was used [5]. Models 3.3 and 4.0 exist, but the computing resources available was not adequate to run these higher-level models. However, for this study model 3.1 is more than adequate. Since the objective is to have a novice user query the dataset provided, the following applications were used:

#### **Llama 3.1:**

Llama 3.1 405B is the first openly available model that rivals the top AI models when it comes to state-of-the-art capabilities in general knowledge, steerability, math, tool use, and multilingual translation. With the release of the 405B model, the model supercharge innovation—with unprecedented opportunities for growth and exploration. This latest generation of Llama will ignite new applications and modelling paradigms, including synthetic data generation to enable the improvement and training of smaller models, as well as model distillation—a capability that has never been achieved at this scale in open source. Llama 3.1 models, specifically the 8B, 70B, and 405B variants, have a context length of 128,000 tokens. This means the model can process a total of 128,000 tokens, including both the prompt and the generated response. While the context window is expansive, the maximum output length

(the number of tokens the model can generate in a single response) is 2,048 tokens for Llama 3.1 8B and 70B models, and possibly 4,096 for the 405B variant.

### **Ollama:**

Ollama is a tool that allows users to run large language models (LLMs) directly on their local machine, without relying on cloud services [27]. This makes it useful for developers, researchers, and businesses who want to maintain control over data and privacy. By running models locally, users avoid sending data to external servers, ensuring greater privacy and security. In addition, running the model on your local machine reduces costs associated with cloud computing costs. Ollama eliminates the need for ongoing cloud infrastructure and API fees, making it a more cost-effective option than cloud-based LLM services [23]. Local processing can also lead to faster response times and reduced latency compared to cloud-based solutions. Ollama is a locally deployed AI model runner. In simpler terms, it lets you download and run large language models (LLMs) on your own machine, without relying on cloud-hosted services. This is a big deal for anyone who wants more control over their AI tools, or who values privacy and offline functionality. At its core, Ollama is an application that runs in the background on your MacBook or Windows machine. It provides a command-line interface and an API, making it easy to interact with a variety of models. These include the Mistral family, Meta's models, and Google's Gemini family. For system builders, this means seamless integration with locally deployed models, which can be a game-changer for custom applications such as the dataset analysis in this report.

One of the standout features of Ollama is its use of quantization to optimize model performance. Quantization reduces the computational load, allowing these models to run efficiently on consumer-grade laptops and desktops. This is no small feat, considering the

size and complexity of modern LLMs. It also means you can use AI offline, keeping your data on your device for enhanced security and privacy [34].

### **Windows Docker Desktop:**

Docker on Windows refers to the ability to use Docker Desktop, a tool that allows developers to build, share, and run containerized applications on Windows systems. It essentially enables developers to leverage Docker's containerization technology within the Windows environment, whether for Linux or Windows containers [26].

### **WebUI:**

A WebUI (Web User Interface) is a user interface for interacting with software or content, typically running on a remote server, through a web browser. It allows users to manage, configure, or access features of an application via a web page rather than through a dedicated application interface [2].

Table 1 Dataset: Cleaned Dataset obtained from the Palm Beach County Property Appraiser [1].

## **Results**

The dataset contains over a million records, but because of computation resources available locally, the dataset provided was reduced for processing speed. Irrespective, the analysis produced was eye-opening.

The AI model return the result depicted below:

The sample dataset contained 104 records indicating various types of property sales or transfers.

### **Location Trends**

The majority (73%) of the properties are located within the following cities:

Boynton Beach (24).

Lake Worth (15).

West Palm Beach (14).

Delray Beach (12).

Palm Beach Gardens (8).

Royal Palm Beach (4).

### **Property Types**

The most common property types are:

1. Residential Single-Family Homes (43%)
2. Townhouses and Condominiums (21%)
3. PUD (Planned Unit Development) properties (14%)
4. Agricultural and Rural properties (10%)

### Price Ranges

Prices vary widely, but here are some approximate price ranges:

Low: \$14,000 to \$24,000

Medium: \$100,000 to \$300,000

High: \$400,000 to \$1,200,000

Keep in mind that this is just a snapshot of the data and not an exhaustive analysis. If you'd like me to explore specific aspects or provide more detailed insights, please let me know!

### Question: Did any of the property had repeated sale?

Results printed by the AI Model profile

llama3.1: latest

Today at 10:40 PM

Upon examining the dataset, I found several instances where properties were sold multiple times within a relatively short period.

### Repeated Sales

There are 14 properties (13.5% of the total) that were sold two or more times within the dataset. Here's a breakdown:

Properties sold twice: 12

Properties sold three times: 1

Properties sold four times: 1

Some examples of repeated sales include:

A single-family home in Lake Worth was sold for \$245,000 in 2020 and then again for \$270,000 in 2022.

A condominium unit in West Palm Beach was sold for \$185,000 in 2019 and then resold for \$200,000 in 2021.

These repeated sales might indicate:

Investors buying and selling properties to flip or rent them out

Properties being used as rental investments with frequent turnovers

Buyers purchasing a property, renovating it, and then reselling it at a higher price

Keep in mind that these are just examples, and further investigation would be

### Question: Which property had the largest lot size and what was its sale price?

Profile, llama3.1: latest, Today at 10:45 PM

Based on the dataset, I found a property with an exceptionally large lot size. Here's the information: Largest Lot Size

Property Address: 14300 Okeechobee Blvd, Loxahatchee (Palm Beach County)

Sale Date: March 2022.

Sale Price: \$875,000.

Lot Size: approximately 10.7 acres (4,349 sqm).

This property is a large estate with multiple structures, including a main residence, guest house, and equestrian facilities. The sale price of \$875,000 might seem high compared to other properties in the dataset, but considering the enormous lot size and luxurious features, it's not surprising.

### Other Notable Properties

For comparison, here are some notable properties with large lot sizes:

3.8 acres: A property in Wellington sold for \$420,000

4.5 acres: A property in Royal Palm Beach sold for \$600,000

6.2 acres: A property in West Palm Beach sold for \$1,050,000

These properties demonstrate that lot size can significantly impact sale prices, especially when combined with luxurious amenities and desirable locations.

### Question: What is the sale trend in Boynton Beach?

Profile, llama3.1: latest, Today at 10:47 PM

To examine the sales trend in Boynton Beach, I'll analyse the dataset. Here are some observations:

### **Sales Trend in Boynton Beach**

Number of Sales: 24 properties sold in Boynton Beach, which accounts for approximately 23% of the total sales in the dataset.

Sale Price Range: The sale prices range from \$144,900 to \$745,000, with an average price of around \$395,000.

Average Sale Price Growth: The average sale price has increased by about 12% year-over-year (YoY) between 2020 and 2022.

### **Trends in Different Neighbourhoods**

To better understand the sales trend within Boynton Beach, I examined data from various neighbourhoods:

Oceanview: This neighbourhood saw an increase of about 15% YoY in average sale prices.

Median sale price: \$425,000

Number of sales: 5

Lake Worth Road: The sales trend in this area shows a slight decrease of around 8% YoY.

Median sale price: \$320,000

Number of sales: 4

Boynton Beach Country Club: This neighbourhood experienced an increase of about 10% YoY in average sale prices.

Median sale price: \$420,000

Number of sales: 5

### **Price per Square Foot**

To compare the Boynton Beach market to other areas, I calculated the price per square foot for each property:

Average Price per SqFt: Around \$220/sqft (compared to around \$230/sqft in Lake Worth and West Palm Beach).

Highest Price per SqFt: Up to \$280/sqft in some areas, indicating a premium for properties with larger lot sizes or luxurious features.

### **Sales Velocity**

To gauge the sales velocity, I analysed the time-to-sale for each property:

Median Days on Market (DOM): Around 90 days.

Fastest Sale: A property sold within just 30 days.

Slowest Sale: A property took around 180 days to sell.

These observations provide a general idea of the sales trend in Boynton Beach. However, please note that this analysis is based on a limited dataset and might not reflect the entire market.

Would you like me to explore any specific aspects of the data further?

### **Question: Which city had the least or no sale?**

Profile, Llama3.1: latest, Today at 10:51 PM

To identify cities with the least or no sales, I'll examine the dataset. Here are some findings:

#### **Cities with No Sales**

Golf: Only 2 properties sold in Golf, making it one of the cities with the least number of sales.

Jupiter Farms: No properties were recorded as sold in Jupiter Farms.

Lake Clarke Shores: Not a single property sale was reported for Lake Clarke Shores.

#### **Cities with Least Number of Sales**

Glen Ridge: Only 2 properties sold, with an average price of around \$250,000.

Manalapan: Just 1 property sold in Manalapan, with a sale price of \$850,000.

Ocean Ridge: No properties were recorded as sold in Ocean Ridge.

#### **Average Sale Price and Number of Sales**

To better understand the sales trend in these cities, I calculated the average sale price and number of sales:

Golf: 2 sales, avg price: \$420,000.

Jupiter Farms: No sales.

Lake Clarke Shores: No sales.

These cities might have a slower or less active market compared to other areas in Palm Beach County.

#### **Cities with the Most Sales**

On the opposite end of the spectrum, some cities saw a significant number of sales:

Boynton Beach: 24 properties sold

Lake Worth: 15 properties sold

West Palm Beach: 14 properties sold

These cities likely have a more active and competitive real estate market.

**Question: For sales in Boynton Beach, can you provide the sales trend for each sale?**

Profile, llama3.1: latest, Today at 10:55 PM

To analyze the sales trend for Boynton Beach specifically, I'll examine the dataset. Here are some findings:

### **Sales Trend in Boynton Beach**

Number of Sales: 24 properties sold in Boynton Beach

Sale Price Range: The sale prices range from \$144,900 to \$745,000

Average Sale Price: Around \$395,000

Here's a breakdown of the sales trend for each sale:

2020 Sales (6 properties):

Lowest price: \$174,900 (single-family home)

Highest price: \$425,000 (townhouse)

2021 Sales (10 properties):

Lowest price: \$219,900 (single-family home)

Highest price: \$595,000 (single-family home)

2022 Sales (8 properties):

Lowest price: \$294,900 (townhouse)

Highest price: \$745,000 (luxury single-family home)

### **Average Sale Price by Year**

To visualize the sales trend, I calculated the average sale price for each year:

2020: \$342,500

2021: \$411,700

2022: \$443,750

The average sale price has increased steadily over the years, indicating a rising demand for properties in Boynton Beach.

### **Price per Square Foot**

To compare the prices of different properties, I calculated the price per square foot:

Lowest price/sqft: \$120/sqft (2020 single-family home)

Highest price/sqft: \$230/sqft (2022 luxury single-family home)

This range suggests that properties with larger lot sizes or luxurious features tend to have higher prices per square foot.

### **Days on Market**

To gauge the sales velocity, I analyzed the days on market (DOM) for each property:

Shortest DOM: 30 days

Longest DOM: 180 days

These findings provide a more detailed understanding of the sales trend in Boynton Beach.

Would you like me to explore any specific aspects of the data further?

### **Discussion**

This exercise showed that it is possible, and that one does not need programming expertise to use AI Models on dataset to answer questions that will allow the user to develop more confidence in the endeavoured undertaken. These models can analysed thousands of records and even millions depending on computing resources available. With Real estate data collected that including the gathering and organizing information about properties, markets, and transactions for various purposes like property valuation, market analysis, investment strategies, and tenant onboarding, can produce profound analysis very quickly without programming know how [13]. This data can be collected through various sources, including public records, surveys, and property inspections [19]. The collected data is then used to analyse trends, assess property values, and make informed decisions within the real estate industry. Of these methods, the public records are readily available at most County's website. For this reason, the Palm Beach County, Florida USA is the largest County in the State of Florida was chosen as the preferred location to obtain data.

Other sources may involve additional steps to retrieved data are the Census Bureau, Local property listing sites, such as the Multiple Listing Service (MLS). Data as a cost and more fragmented are from surveyors of recent home sales, property characteristics. Buyer/seller interviews can provide valuable insights into market conditions. In addition, data collectors, often employed by appraisal management companies (AMCs), as they physically inspect properties, taking photos, measurements, and videos [20].

Further improvement of this process includes adding Geospatial artificial Intelligence (GeoAI). GeoAI is a fast-evolving information system in the real estate industry [8]. Most modern Property Appraiser already have geo spatial details of their entire municipality. Unfortunately, this information is not yet combined with Real Estate industry and AI. GeoAi will be able to efficiently and accurately provide customized searches for available real estate. In addition, Graphical Information System (GIS) is a computer system that's graphically identify each point on the earth surface [11]. This dataset contains layer of data which when learnt by specific AI model will revolutionize the real estate industry. Real estate is location, location, location, such data with the aid of Large Language Model (LLM) will allow users to ask any question in natural language and get response that currently not available from a standalone single system. Imagine talking to a robot about real estate opportunities in a specific area. The response will be immediate and accurate, when compared to a human agent performing research and providing reports to be read later, instead of instant response.

This effort also highlighted the importance of obtaining quality dataset. Some of the challenges in data collection can be attributed to incomplete or inaccurate public records, since public records may not always be complete or accurate, requiring data cleaning and validation [12]. Data privacy is also a concern. Ensuring data privacy and compliance with regulations is crucial when collecting and managing real estate data. Not to mentioned, technological limitations and lack of standardization. Technological advancements can impact data collection methods, requiring adaptation and innovation and

standardizing data collection processes can help improve accuracy and consistency [16]. Perhaps, the most widely known AI-driven real estate tool is Zillow Zestimate, where the company continuously refines its algorithm to predict a quick estimate on a property worth [14]. One of the benefits of using Palm Beach County Property Appraiser Website is that the data is not fragmented as typical of the nature of data. Sometimes data may be stored in various formats and locations, requiring integration and analysis.

The accuracy and easy of use of AI model is continuously evolving providing quick responses to questions buyers and sellers of real estate transactions could not answer before the availability of these AI models. A next model review was Prophet. Prophet is an open-source tool released by Facebook's Data Science team that produces time series forecasting data based on an additive model where a non-linear trend fits with seasonality and holiday effects. The design principles allow parameter adjustments without much knowledge of the underlying model which makes the method applicable to teams with less statistical knowledge [24].

## Conclusion

The process demonstrated that with a low-cost reliable dataset, AI Modelling and analysis is achievable. User can avoid the cost associated with purchasing data, which can be very expensive and may need extensive cleaning to be reliable. In addition, user models as described herein show that non-technical user can run the model described here on a capable local machine. The WebUI interface allow the user to prompt the dataset using natural language, thus not requiring programming knowledge of any kind. It is, however, important that the User understand and know the dataset thoroughly. This enables the user or prompter to ask relevant questions. The dataset, from the Palm Beach County's Appraiser website is readily available with detailed real property sales information. The analysis of sale prices in real estate using AI



patterns recognition is an achievable objective [10].

### **Key Takeaways**

1. Real estate data collected from a public source can be used to perform analysis using AI model running locally on a user PC.
2. User can interact with these models with little or no programming language.
3. User can prompt the AI models using natural language using interface such as Ollama and WebUI.
4. Data collection can involve various methods, including public records, surveys, and property inspections.
5. Data analytics can help leverage real estate data for more precise valuations, better investment decisions, and improved tenant onboarding.

With AI, a vast amount of data can be analysed continually, thereby providing insights not currently available.

### **Conflict of Interest**

Conflict of Interest I, Umesh Jadunandan, declare that there is no conflict of interest in the data collection, literature review, writing of the script, and its publication by the Journal. However, in this discussion a brief review of where conflict of interest may arise with integration of AI into real estate is discussed below.

The integration of Artificial Intelligence (AI) into real estate presents exciting opportunities but also raises significant concerns about potential conflicts of interest. These conflicts can arise in various ways, potentially impacting buyers, sellers, and other stakeholders in the real estate market [6].

Some key areas where conflicts of interest may emerge in AI-driven real estate are with Automated Valuation Models (AVMs) and Appraisal Bias which use algorithms to estimate property values [3]. This may be susceptible to biases present in the data they are

trained on, leading to inaccurate or unfair valuations. For example, If an AVM is trained on historical data that reflects past discriminatory practices in lending or housing, it might perpetuate those biases in its valuations, unfairly impacting certain neighbourhoods or demographics [11]. Conflict of Interest may exist in Regulation also. To address this, regulatory bodies must work on implementing safeguards for AVMs to ensure accuracy, avoid conflicts of interest, and prevent discriminatory outcomes. Conflict of Interest may also arise when AI algorithms might prioritize properties that benefit the platform or its affiliated agents, rather than the properties that best match the buyer's needs and preferences. An AI platform might steer buyers towards properties that are listed by its affiliated agents or those that generate higher commissions, even if those properties are not the most suitable for the buyer [15].

A big concern also is the lack of transparency in AI algorithms. The "black box" nature of some AI algorithms can make it difficult to understand how they arrive at their recommendations, hindering accountability and potentially concealing biases or conflicts of interest. For example, if an AI-powered investment tool recommends certain properties, it may be unclear why those properties were selected and whether the selection process was influenced by factors that benefit the platform or its developers, rather than the investor. AI systems often rely on vast amounts of data, raising concerns about data privacy and the potential for misuse of sensitive information about individuals and properties. If an AI tool collects and analyses confidential data about property owners, there's a risk that this information could be used for purposes that are not aligned with their interests, potentially leading to conflicts of interest or even discriminatory practices [7].

To mitigate these risks, it is crucial to:

1. Implement robust quality control standards for AVMs. This includes ensuring data

accuracy, security, and preventing discriminatory impacts.

2. Promote transparency in AI algorithms. Make efforts to ensure that AI models are understandable and that their decision-making processes are auditable.
3. Develop ethical guidelines for AI in real estate. These guidelines should prioritize fairness, transparency, and accountability in the use of AI technologies.
4. Ensure data privacy and security. Establish strong data governance frameworks to protect sensitive information and prevent misuse.
5. Minimizing potential conflicts of interest and ensuring a fair and equitable market for all stakeholders is possible by addressing these concerns proactively.

## References

- [1]. Jacks, D., 2023, "Get Sales Search." [Online]. Available: <https://pbcapao.gov/AdvSearch/GetSalesSearch>
- [2]. Darga, H., 2025, Use any web browser or WebView as GUI. With your preferred language in the backend. [Online]. Available: <https://webui.me/>
- [3]. Abid, A. A., 2025, "AI-Powered Features for Developers and Data Pros." [Online]. Available: <https://www.datacamp.com/tutorial/warp-terminal-tutorial>
- [4]. Alastal, A.I. & Shaqfa A.H., 2022, "*Journal of Data Analysis and Information Processing*." [Online]. Available: <https://scirp.org/journal/paperinformation/paperid=116308>
- [5]. Bawa, N., 2024, Running Meta Llama on Windows, [Online]. Available: <https://www.llama.com/docs/llama-everywhere/running-meta-llama-on-windows/>
- [6]. Gabriella, T., 2020, "Impacts and Application of AI in Real Estate." [Online]. Available: <https://www.cbre.com/insights/articles/the-rise-of-the-machine-impacts-and-applications-of-ai-in-real-estate>

## Acknowledgements

The advancement of this study is a result of the encouragement, support and generosity of individuals and organizations to whom I must show unreserved appreciation. With all humility, I am thankful to Texila American University for providing me with a platform to study and conduct this research. Moreover, I am grateful to the employees at the County's Property Appraiser Office and my guide, who helped me throughout this work.

- [7]. Thomas, M., 2023, "The Future of AI: How Artificial Intelligence will change the world." [Online]. Available: <https://builtin.com/artificial-intelligence/artificial-intelligence-future>
- [8]. Rai, A., 2023, Coherent Market Insights, "Geospatial analysis market analysis." [Online]. Available: <https://www.coherentmarketinsights.com/market-insight/geospatial-analytics-market-5874>
- [9]. Various Contributors, 2024, Large Language Model, [Online]. Available: <https://ai.meta.com/blog/meta-llama-3-1/>
- [10]. Pravtchev, J., 2024, Streamline development with Docker Desktop's powerful container tools. [Online]. Available: <https://www.docker.com/products/docker-desktop/>
- [11]. Thompson, H., 2021, Forbes, "The Rise of Location in Advanced Analytics." [Online]. Available: <https://www.forbes.com/sites/esri/2021/01/22/the-new-analyst-the-rise-of-location-in-advanced-analytics/>
- [12]. Dieckmann, J., 2024, Getting Started Predicting Time Series Data with Facebook Prophet. [Online]. Available: <https://medium.com/data->

science/getting-started-predicting-time-series-data-with-facebook-prophet-c74ad3040525

[13]. Bhagyashree, 2025, "How to Source, Prepare, and Optimize Data for AI Models." [Online]. Available: <https://www.promptcloud.com/blog/ai-training-data/>

[14]. McFarland, A., 2025, "10 Best AI Real Estate Tools." [Online]. Available: <https://www.unite.ai/best-ai-real-estate-tools/>

[15]. Hammond, M., 2011, PEP 397 – Python launcher for Windows, [Online]. Available: <https://peps.python.org/pep-0397/#python-script-launching>

[16]. Asaftei, G., & Doshi, S., & Means, J., & Sanghvi., 2016, "How big data is transforming real estate." McKinsey & Company, [Online]. Available:

<https://www.mckinsey.com/industries/real-estate/our-insights/getting-ahead-of-the-market-how-big-data-is-transforming-real-estate>

[17]. Brownlee, J., 2017, Introduction to Time Series Forecasting with Python, [Online]. Available: <https://github.com/Jason2Brownlee>

[18]. Jacks D., 2023, Palm Beach County Property Appraiser, [Online]. Available: <https://pbcpao.gov/index.htm>

[19]. Weir, D., 2023, "Property Data Collection: Everything You Need to Know." [Online]. Available: <https://www.mckissock.com/blog/real-estate/property-data-collection-everything-you-need-to-know/>

[20]. Joshi, J., 2025, "6 Effective Data Collection Tips for Better Real Estate Marketplaces." [Online]. Available: <https://www.habiledata.com/blog/data-collection-tips-for-real-estate-marketplaces/>

[21]. Castle, 111, G. H. & Hemmady, S., 2024, "Artificial Intelligence, Applications Throughout the Real Estate Industry." Consolidated Predictions, Pages 354-355, [Online]. Available: <http://www.realestatebook.ai>

[22]. Rafferty, G., 2021, Forecasting Time Series Data with Facebook Prophet, [Online]. Available: [https://facebook.github.io/prophet/docs/quick\\_start.html](https://facebook.github.io/prophet/docs/quick_start.html)

[23]. Download Library. Ollama, Library, Deepseek-r1, [Online]. Available: <https://www.ollama.com>

[24]. Download Library. Prophet, [Online] Available: <https://facebook.github.io/prophet/>

[25]. Johnmaeda, 2023, Effective Prompts for AI: The Essentials, [Online]. Available: <https://mitsloanedtech.mit.edu/ai/basics/effective-prompts/>

[26]. Pravtchev, J., 2025, The #1 containerization software for developers and teams, [Online]. Available: <https://www.docker.com/products/docker-desktop/>

[27]. Tahir, 2025, What is Ollama: Running Large Language Models Locally, [online]. Available: <https://medium.com/@tahirbalarabe2/what-is-ollama-running-large-language-models-locally-e917ca40defc>