ZoningLLM – A Novel Multimodal Application for Zoning Analysis

Simha Kalimipalli
University of Waterloo
nkalimip@uwaterloo.ca

Saurodeep Majumdar
University of Waterloo
sdmajumdar@uwaterloo.ca

Kevin Tan
University of Waterloo
kx3tan@uwaterloo.ca

Jonathan Feng
University of Waterloo
j25feng@uwaterloo.ca

Rahul Kumar
University of Waterloo
r77kumar@uwaterloo.ca

Liam Dachner
University of Waterloo
ldachner@uwaterloo.ca

***Abstract*—Large municipalities in Canada have recently faced an unprecedented housing crisis. This has been driven by an increase in the demand for housing and a lack of housing supply. Stringent zoning requirements have contributed to reducing the construction of new housing. Each municipality typically has its own separate zoning code consisting of lengthy documents written in technical jargon. It is difficult for the public, researchers, and home builders alike to extract relevant information from these documents. This opacity restricts the discussion of zoning policy and aggravates the housing crisis. This project aimed to use generative and geomatic AI methods to analyze zoning and construction documents for the Waterloo Region to gain insights about zoning restrictions. This can be used to quantify and monitor the effects of zoning on housing supply. A web-based application with the ability to process, export and query ad-hoc zoning queries has been developed. Discussions with regional planners have underscored the importance of this work.**
Keywords: Large-Language Models, Zoning, Housing

I. Introduction

A. Background

Zoning regulations for a municipality typically consist of several main zoning districts and respective subdistricts [1]. Maximum and minimum requirements for each zoning district are typically listed. Additional, special clauses are often stipulated for special overlapping districts, grandfathered lots and other miscellaneous information. Zoning conditions are reflective of current conditions of a given plot of land rather than development priorities. Official plans are a document, which reflect municipal intensification and urban growth priorities. They are often less structured than Zoning regulations and reflect the long-term vision of the municipality. Both zoning regulations and Official Plans are important for assessing future housing supply [2].

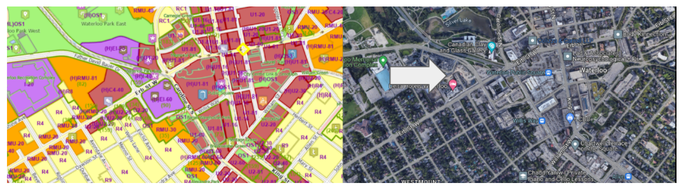

Fig. 1. Uptown Waterloo Zoning: redevelopments (at arrow) and zoning constrained neigbourhoods (R4) are pictured

B. User Needs Research

We conducted research to assess important factors for zoning codes and better understand the development process. We spoke to current urban planners at the Region of Waterloo, a development lawyer in British Columbia, and other key stakeholders related to zoning.

We researched current procedures by 3 stakeholders (developers, researchers and policymakers) which could be improved by a zoning related application:

1) Developers

They follow 5 steps in development:

1) **Find macro purpose:** They see where there is a mismatch between official plan (long term vision), zoning codes (current conditions) and aggregated demand to identify potential for housing/development.
2) **mid-level opportunity identification:** Developers evaluate initial plans based on criteria, such as financing, location and granular demand.
3) **Micro-level Analysis:** They figure out if local infrastructure (transportation, electrical, waste water etc) supports rezoning. City council attitudes towards previous house development is also an important consideration. This includes future plans/approvals data and the presence of density bonusing incentive programs.
4) **Comprise for micro level decisions:** Negotiations with Policy-makers and local groups occur until a compromise is reached, and compliance is met. The developer typically aims for either highly profitable or specialty "landmark" projects.

5) **Commencement of Construction** Actual construction of buildings occur.

2) Researchers

They currently manually parse through zoning codes to extract useful information. This is often a manual and slow process. They are interested in correlation of these zoning code data with development patterns. They need to quickly and systematically collect zoning data for extensive analysis to be feasible.

3) Policy Makers

They want to increase transparency of open data, investing in new useful tools to help with that goal. They want to make better policy decisions (eg. seeing where official plans/zoning codes are adequate/inadequate and changing zoning). They have to negotiate with developers and community local groups to reach a compromise on final building plans.

C. Related Works

It is important to assess the relevant literature regarding scraping information from zoning regulations:

1) Academic Literature

We were inspired by the Urban Institute's work on Automating Zoning Data Collection. Their seminal paper discusses creating a unified zoning database for manual scraping of data from the zoning districts within municipalities. LLMs gain traction for decoding regulatory information across various domains, such as the financial and health sectors [3].Barthik, Gupta et al discuss using Large Language Models to decode zoning statutes in the United States. [4]. This paper also discusses creating a unified database for zoning regulatory information, but the data granularity is at the municipality level (where each municipality is assessed as a whole) unlike the Urban Institute's work.

2) Industry Efforts

There are also several startups doing relevant work: Trax.co is a startup that is working on creating LLM accessible building codes in Ontario [5]. Arterial.design is a Boston based startup that "automates decision-making for policy-driven organizations" [6]. Up.codes is a platform to streamline code compliance for architects, homebuilders and inspectors [7]. Finally, Autoprop is a Vancouver-based company that provides data automation solutions for real estate professionals. These companies demonstrate the market viability of using LLMs, web mapping software to create databases that simplify regulatory information [8].

3) Literature Conclusions

Based on the literature, relevant companies and interviews conducted, our design project aims to be the first project to use multimodal LLMs to automatically retrieve zoning data by zoning districts and visualize and analysis this data through a geomatic lens.

II. Methodology

A. Problem Definition

The team proposes using NLP methodologies such as Retrieval Augmented Generation to 1) simplify the verbosity of various zoning documents to illustrate what can and cannot be built in a certain area (ie. by chatbot). 2) Extract exact details from zoning documents about setback and density. 3) Use classical ML techniques to correlate extracted zoning information with housing development in a given municipality. 5) Create a unified framework to compare the zoning of neighborhoods across the 3 Urban municipalities in the Region of Waterloo.

The ultimate goal of the application would be for the general public to have increased awareness of zoning laws, while providing a useful resource for researchers and policy makers.

B. Dataset Design

1) Phase 1

We then collected zoning regulations for the 3 Urban municipalities of Waterloo Region that are available as PDFs on the municipalities' respective "Zoning and Building" pages.For example, Kitchener's zoning data is found on their Open Data site [9]. This formed the backbone of our data repository. We then exported this data into PDFs which would be later be fed to the LLM [10].

2) Phase 2 and beyond

The geographical areas of zoning codes in the region were owned by a private organization. Hence, we had to manually trace sample urban districts using Google Maps. We converted this data to GeoJSON and would later combine it with the zoning data for each respective district. We would corroborate addition housing information sourced from Zillow and Waterloo Region Connected . Geographical datasets that Correspond to transportation, energy utility and Water infrastructure were also exported from municipal open data sources.

C. Technology used

We strived to identify high-accuracy, cost–efficient tools and technologies for the project.

LLM + RAG Based projects typically have raw textual data that is sent into a vector database, which first turns textual meanings into vectorial representations. and stores the relevant information. Afterwards this data is sent as context to a foundational Large Language Model, that aims to find the most similar internal source data to a prompt [11].

Front End Frameworks are required for the GUI of the application. The front end would either use custom made or ready-made JavaScript/HTML/CSS rendering templates (ie. through Streamlit or Azure).

Back End Frameworks are used logic of the application. Most of the backend of this project would be done in the Python language in conjunction with various LLM APIs. We initial hoped to evaluate the use of several models such as GPT, Claude, Gemini and LLAMA for the LLM similar to other projects.We are evaluating the use of Langchain, and the "Agentic" RAG options for the various LLMs. We are also looking at cloud computing/deployment services such as Heroku, Amazon Web Services or Azure to deploy to application when completed.

III. Results

A. Phase1 - Base Tested LLM

1) Iteration1

The first iteration used the Azure CosmosDB Vector Database, GPT 3.5 Turbo as the LLM and had an Azure studio hosted webapp for the Interface (Backend + GUI).

We noticed that while this version was easy to set up and worked decently for small question, The performance was degraded for more complicated queries (circular and non answers were common). Additionally model hyper-parameters were not able to be extracted. Hence a second iteration was explored.

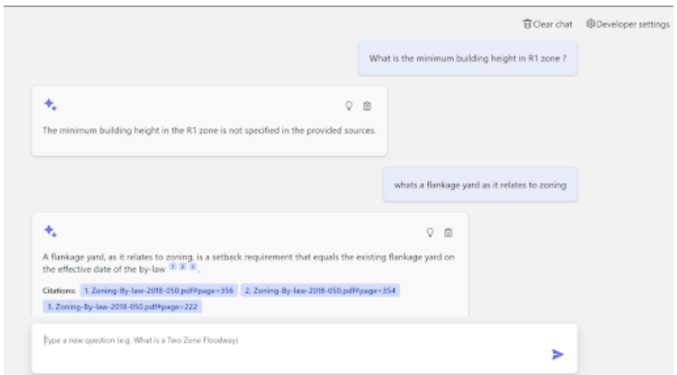

Fig. 2. Chatbot with Testing Questions

2) Iteration2

A second iteration of the product was also made after further research. This version uses the FAISS Database, GPT 4 Turbo LLM, Streamlit for the front end and Langchain for orchestration. This version is currently locally hosted. Testing of the second iteration also included the same sample questions had the first one but also included some real hyperparameters: model, chunk size, chunk overlap and temperature. The model is adjustable using GPT 3.5 Turbo GPT 4 Turbo currently (other models undergoing testing).

Chunk size refers to the largest character size of a text chunk used for embedding. Chunk overlap is the number of characters that are overlapped between two chunks that are next to each other for vectorization. Temperature refers to the amount of randomness of chatbot results. [12].

TABLE I
Hyperparameter Options and Limits.

Hyperparameter	Tuning limits
Chunk size	[100,2000]
Chunk Overlap	[10,200]
Temperature	[0.05,0.3]
Zoning Codes	[R1,R4,RMU]

A set of 13 Questions were prompted to the LLM which were sourced from the Barthik and Gupta paper, the National Zoning Atlas paper, and advice from current urban planner (s) [4] [10] .

1) What is the minimum building height?
2) What is the maximum building height?
3) What is the minimum street line setback?
4) What is the maximum street line setback?
5) What is the minimum density?
6) What is the maximum density?
7) What are minimum frontage requirements for single family residential development?
8) What are maximum frontage requirements for single family residential development?
9) Are apartments above commercial (mixed use) allowed?
10) Is multi-family housing allowed, either by right or special permit (including through overlays or cluster zoning)?
11) Are attached single family houses (townhouses, 3+ units) listed as an allowed use (by right or special permit)?
12) Are accessory dwelling Units (ADUs) or in-law apartments allowed (by right or special permit)?
13) Is cluster development, planned unit development, Planned Residential Development (PRD) open space residential design, or another type of flexible zoning allowed by right?

We customized questions by attaching the word ”in” and one of 3 zoning codes, Residential1, Residential4, and Residential Mixed-use zone to the end of questions. Additionally, we ran each configuration at least 5 times due to the stochasticity of LLMs (which could output different result each time) to ensure reliability.

TABLE II
Best Parameters and Values for Different Zones.

Zone	(Chunk Size, Chunk Overlap, Model, Temperature)	Best Accuracy
R1	(800, 100, gpt-3.5-turbo, 0.26)	1.0
R4 - SINGLE DETACHED	(1000, 100, gpt-3.5-turbo, 0.12)	0.67
R4 - SEMI-DETACHED and DUPLEX	(800, 100, gpt-4-turbo, 0.35)	0.83
R4 - FREEHOLD SEMI DETACHED	(600, 100, gpt-4-turbo, 0.012)	0.83
RMU-20	(800, 100, gpt-4-turbo, 0.10)	0.83
RMU-30	(1000, 100, gpt-4-turbo, 0.05)	0.83
RMU-40	(600, 100, gpt-4-turbo, 0.49)	1.0

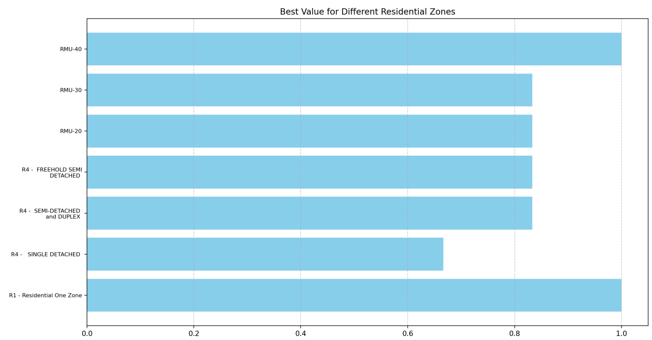

Fig. 3. Accuracy over different zones

Results indicate that the model is able to perform accurately across different zoning and sub-zoning districts. The higher performance is shown with GPT-4-turbo model using lower value of temperature. The model is able to do considerably well even in zones that have multiple subzones (ie. RMU). This indicates a high degree of ability for the model to a) read and comprehend tables with requirements for multiple subzones as well as b)retrieve the relevant content from its database to be able to answer the user queries. It can be seen that a moderate chunk size between 600-800 tokens is sufficient to allow for retrieval of relevant content. "R-4 - Single Detached" was the only zoning code with significant retrieval issues.

Future Phases of the project use the average of the best recall hyperparameters for each zoning code identified in Phase 1.

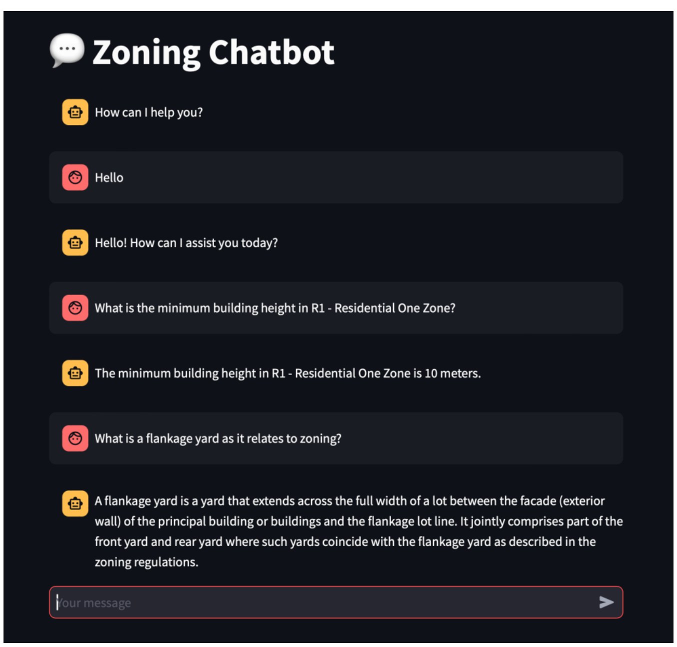

Fig. 4. Second Iteration of Chatbot with testing questions

B. Phase2 - Query and Download

For this phase, Given a suitable query (with municipality, zoning code, sub-zoning code and chosen metrics), the system would output the requested information in a structured format (i.e. in JSON or CSV format) along with an AI-generated summary of zone. The data would come mostly from the zoning code and official plan(undergoing tests).

An example of such a query would be:

```
Municipality: [Waterloo];
Zoning codes: [Residential3, Residential5];
Zub-zoning codes: [All];
Metrics: [maximum height, minimum height,
maximum density, minimum density]
```

We developed random test questions (with known answers) to test the system. Final Output types after prompt engineering are shown below:


```
[
    {
        "Municipality": "City of Waterloo",
        "Zoning": "R1 - Residential One Zone",
        "Metrics": [
            { "Metric": "Lot area", "Unit": "square metres", "Requirement": "Interior Lot: >= 405, Corner Lot: >= 540" },
            { "Metric": "Lot frontage", "Unit": "metres", "Requirement": "Interior Lot: >= 13.5, Corner Lot: >= 18" },
            { "Metric": "Front yard setback", "Unit": "metres", "Requirement": ">= 7.5" },
            { "Metric": "Flankage yard setback", "Unit": "metres", "Requirement": ">= 6" },
            { "Metric": "Side yard setback", "Unit": "metres", "Requirement": ">= 1.8" },
            { "Metric": "Rear yard setback", "Unit": "metres", "Requirement": ">= 7.5" },
            { "Metric": "Building height", "Unit": "metres", "Requirement": "<= 10" },
            { "Metric": "Lot coverage", "Unit": "%", "Requirement": "<= 35" },
            { "Metric": "Parking spaces", "Unit": "per dwelling unit", "Requirement": ">= 1" },
            { "Metric": "Number of buildings", "Unit": "count", "Requirement": "1 main building per lot" }
        ]
    },
```


Fig. 5. Sample JSON output of Query and Download Feature

C. Phase3 - Select and View

Geographic Information pertaining to zoning code districts and Future growth areas was aquired from municipalities in a GeojSON format. We merged our textual zoning data output for each municipality (from the Query and Download Feature) and our geographic data to create a single layer. We would display this information as an interactive layer in the zoning LLM app. We made a system such that queries from our chat bot pertaining to geographical information would be converted to SQL and all districts meeting the query criteria would be selected. A sample query would be "Select all areas with a maximum height of 10" meters". We cross referenced results with the actual zoning map to ensure accuracy of this feature. We used the MapBOX mapping API for this and future phases with mapping.

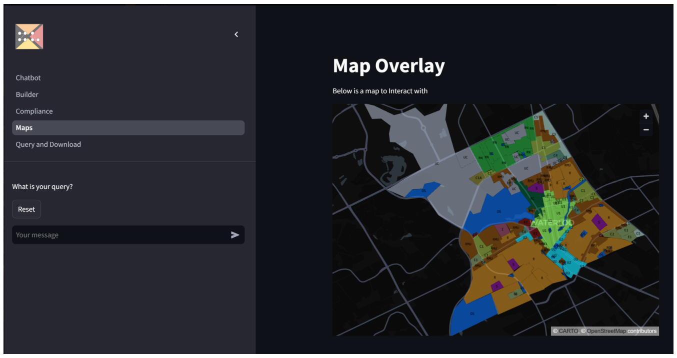

Fig. 6. Map of Zoning Codes in Central Waterloo

D. Phase4 - Historical Data

In additional to the zoning information recall identified in previous phases, We would also add housing supply and demand information from various sources to our mapping/chatbot database. We connect to the APIS of real estate websites such as Zillow for information about units on the market, and Local development forums such as UrbanToronto, or Waterloo RegionConnected which keep detailed databases of recent construction products with information about the size, builders, and number of units [13].

This information could geographically either be in a "point format" (real estate, development information), where each point would represent a unit/development respectively; or in a "area" format for broader neighborhood level insights.

Boolean search operations (NOT, AND, OR) could be used to identify interaction zones between various zoning and housing data. By comparing the various intersection zones (restrictive zones/places with lots of development areas) we can could quantify correlations between zoning regulation and housing development. However, current testing has been restricted to qualitative analysis for the time being.

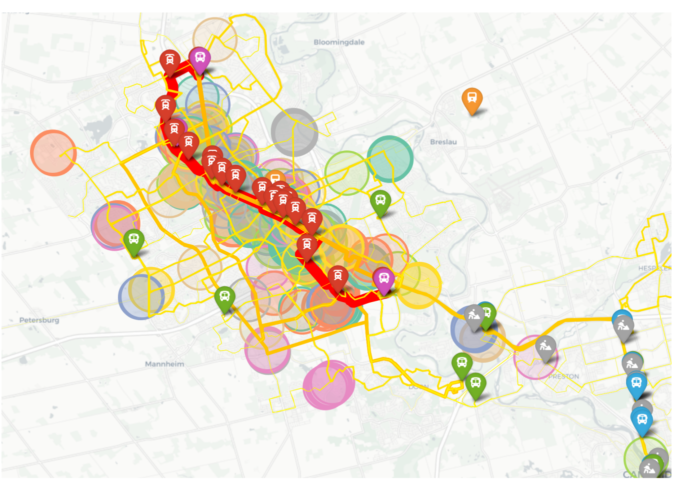

Fig. 7. Map of recent multi-story developments and transport infrastructure in Waterloo

E. Phase5 - Development Compliance and Visualization

Once a plot is selected for development, developers need to construct an example of massing for a proposal. This is a time consuming process that requires constructing detailed 3D models of a proposed structures. These models need to be remade regularly as the development concept changes. Thus, we have developed a system that is able to generate renders and 3D models of massing structure. The user would input the height (in storeys) of the building, the shape (L shape, rectangular etc.), the location (suburban or urban) and a model of the structure is generated using the MapBOX mapping library. Given the maximum building sizes/densities in a zoning code, the application says if the building does not meet the zoning code and it suggests similar but feasible structures.

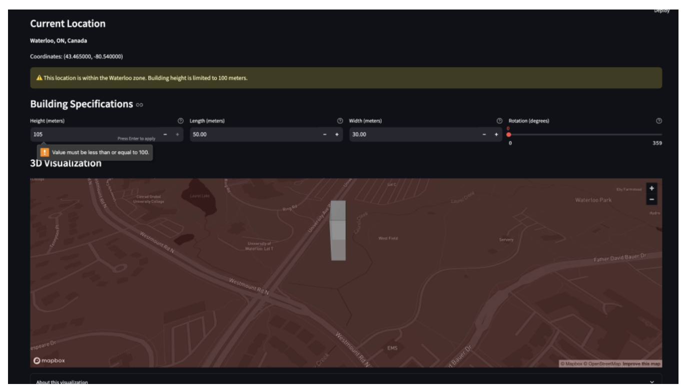

Fig. 8. Sample compliance feature notification

The OpenAI (DALL-E) API is used to create a render of the requested building. Users have the option to download their generated media.

Fig. 9. Sample rendered building

IV. Conclusion

V. Key Findings

The housing crisis is a prominent issue in Canada [14]. Zoning codes are one of the factors which limit housing supply. Generative models have been shown to be effective at extracting relevant information from regulatory documents. This project has applied these models to extract related information from zoning code documents to quantify effects of housing regulation. Extensive benchmarking undertaken during this project has demonstrated that these models indeed show promise in extracting textual information and generating useful insights in geographical, image and textual domains. Researchers and relevant individuals who were contacted during the research process have also shown interest in this technology. Thus, the team will continue to develop this model for enterprise purposes in the future. Specifically features to further analyze historical housing and infrastructure data will be pursued.

VI. Limitations

Hyperparameter Testing only used GPT-3.5 and GPT-4 as LLMs due to budget constraints. Several features were not able to be implemented due to ongoing difficulties in obtaining data (eg. real time rent prices). Although various chunking sizes were used in the RAG process, the chunking process used simple text splitting to extract text from tables. In the future, usage of OCR and object detection techniques can lead to higher accuracy and precision metrics from the model.

VII. Acknowledgments

We would like to thank WAT.ai for their helpful resources. We would also like to acknowledge Sustainable Waterloo Innovation Lab for their efforts in bringing this app to reality. We would finally like to thank Y-Combinator startup Hamming.ai for their advice on LLM benchmarking techniques.

References

[1] K. P. Green and J. Filipowicz, “Barriers to housing supply in ontario and the greater toronto area,” 2023. [Online]. Available: https://www.fraserinstitute.org/sites/default/files/barriers-to-housing-supply-in-ontario-and-the-greater-toronto-area.pdf

[2] X. Gabaix, “Zipf’s law for cities: An explanation,” National Bureau of Economic Research, Working Paper 8835, 2002. [Online]. Available: https://www.nber.org/system/files/working_papers/w8835/w8835.pdf

[3] Z. Cao and Z. Feinstein, “Large language model in financial regulatory interpretation,” Jul 2024. [Online]. Available: http://arxiv.org/abs/2405.06808

[4] A. Bartik, A. Gupta, and D. Milo, “The costs of housing regulation: Evidence from generative regulatory measurement,” *SSRN Electronic Journal*, 2023.

[5] Tra.co. [Online]. Available: https://www.trax.co/

[6] R. Kothari, “Automated land development.” [Online]. Available: https://www.arterial.design/

[7] Upcodes. [Online]. Available: https://up.codes/

[8] AutoProp. [Online]. Available: https://autoprop.ca/

[9] C. of Kitchener, “Open data,” accessed: 2025-03-17. [Online]. Available: https://www.kitchener.ca/en/council-and-city-administration/open-data.aspx

[10] J. Axelrod, L. Lo, and S. C. Bronin, “Automating zoning data collection,” Urban Institute and Cornell University, Tech. Rep., feb 2023, accessed: 2025-03-17. [Online]. Available: https://www.urban.org/sites/default/files/2023-02/Automating%20Zoning%20Data%20Collection.pdf

[11] K. Pandya and M. Holia, “Automating customer service using langchain: Building custom open-source gpt chatbot for organizations,” 2023. [Online]. Available: https://arxiv.org/abs/2310.05421

[12] D. Lin, “Revolutionizing retrieval-augmented generation with enhanced pdf structure recognition,” 2024. [Online]. Available: https://arxiv.org/abs/2401.12599

[13] zillow. [Online]. Available: https://www.zillow.com/

[14] K. P. Green and J. Filipowicz, “The impact of land-use regulation on housing supply in canada,” 2024. [Online]. Available: https://www.fraserinstitute.org/sites/default/files/impact-of-land-use-regulation-on-housing-supply-in-canada.pdf