Building a Multimodal LLM-based Search Assistant Chatbot to Enhance Housing Search

September 2024

Grupo QuintoAndar

Building a Multimodal LLM-based Search Assistant Chatbot to Enhance Housing Search

September 2024

Grupo QuintoAndar

Agenda

1. Who We Are
2. Why a Search Assistant?
3. What it is
4. The Chat Component
5. The Search Component
6. Final Remarks

Who are we?

We are QuintoAndar Group, the **largest real estate ecosystem in Latin America.**

Driven by a shared purpose of helping people love where they live, we offer a diverse portfolio of brands and solutions that cover all stages of the living journey.

We develop technologies and innovations to transform and enhance the housing experience as a whole.

Grupo QuintoAndar

Where are we?

Headquartered in Brazil and with a presence in five other Latin American countries (in addition to a tech hub in Europe), we are a global group that transcends borders.

The QuintoAndar Group currently operates in the following countries: **Brazil, Argentina, Ecuador, Panama, Peru, and Mexico.**

Our team is made up of talented professionals who work from various locations around the world.

Grupo QuintoAndar

Our Numbers

We have formed the largest real estate group in Latin America, and the numbers we've achieved help to illustrate the scale of our brand.

Grupo QuintoAndar

Traffic

45 million monthly visits to our platforms

Listings

Over 8 million property listings published

Assets

$30 billion USD in assets under management

Administration

Over 270,000 contracts under administration

What is the Product

Chat

Real estate agent would talk to you until they **understand what you need...**

So does Copilot!

Search

Once a real estate agent understands your need, they will **find** the property **tailored** to your needs, prioritizing things most important for **you...**

Similar as Copilot!

Why did we need LLM-based Search Assistant?

Before

Does apartment have wooden floor?

Need to review all photos manually

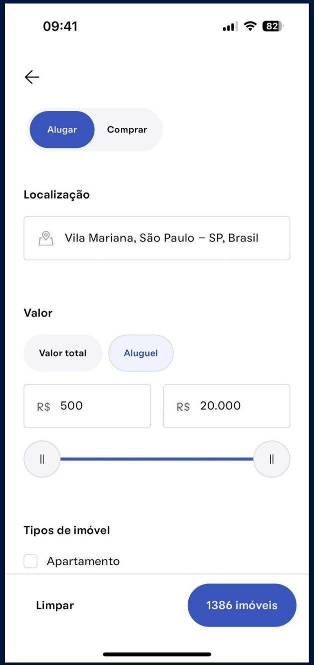
Is apartment close to supermarkets?

Need to go to the map and check

Me and my husband have 2 kids, we need to find nice apartment to fit us all close to the good schools

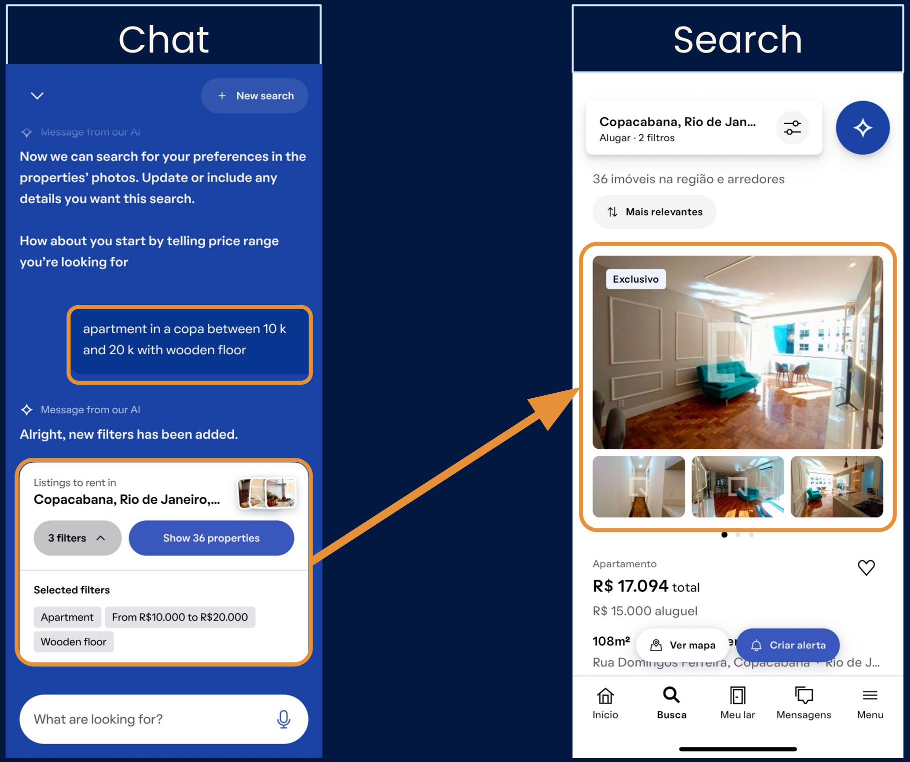
Need to adapt requirements to the filter structure

Search queries are filter-based only



Why did we need LLM-based Search Assistant?

Now



Why did we need LLM-based Search Assistant?

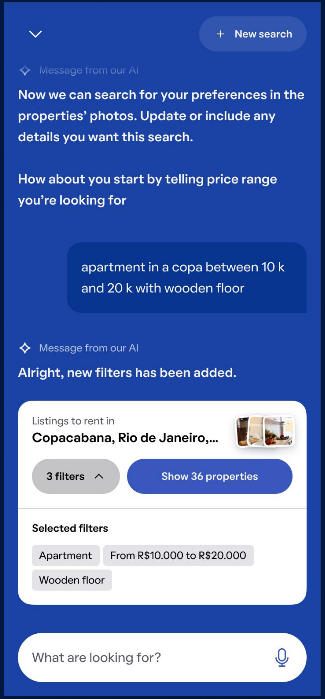
Now

1. ### A Data Collector Chat

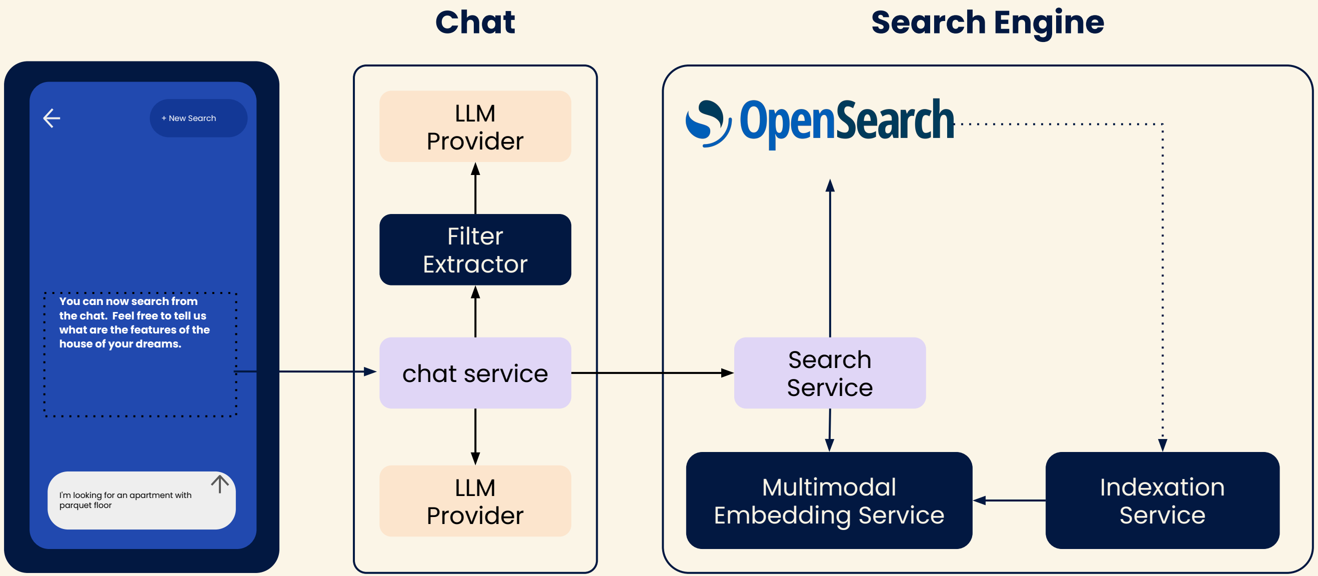
A chat that prompts the user to provide preferences until enough data is gathered to perform a search

2. ### A Search Engine

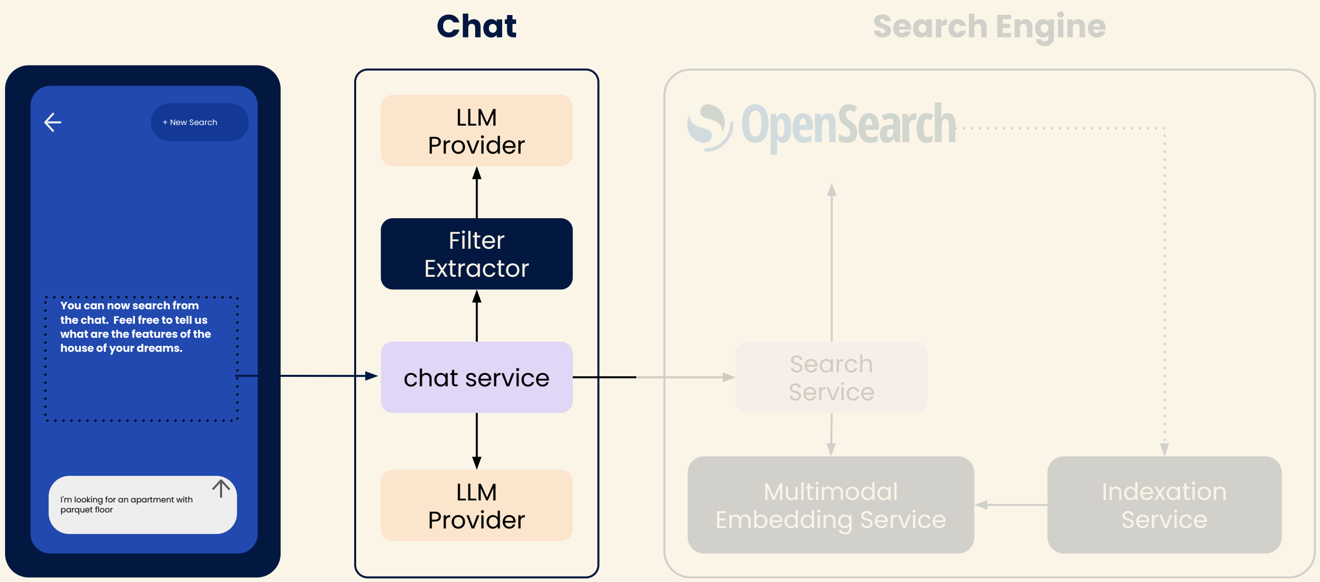
A tool that converts the user's preferences, including visual preferences, into a search.



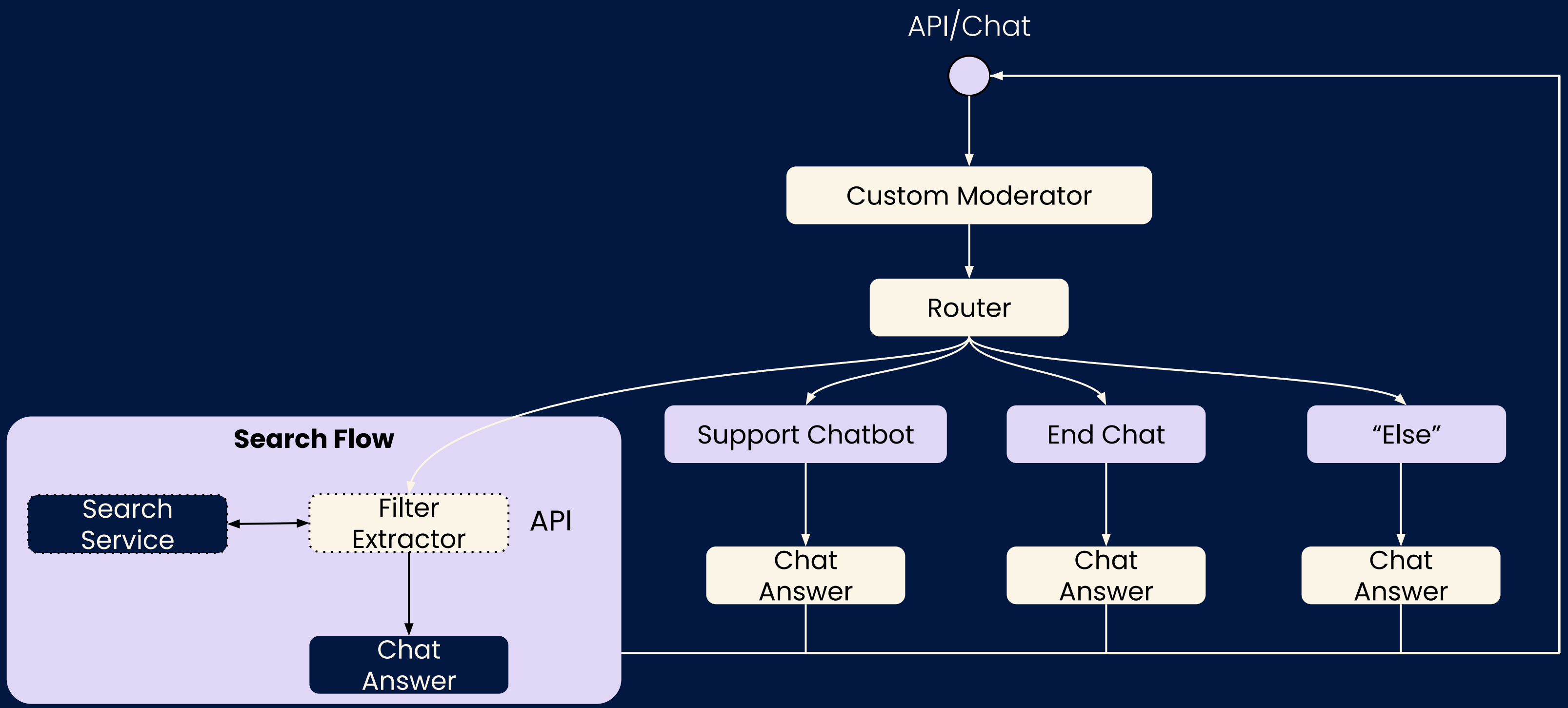
LLM-based Search Assistant



LLM-based Search Assistant

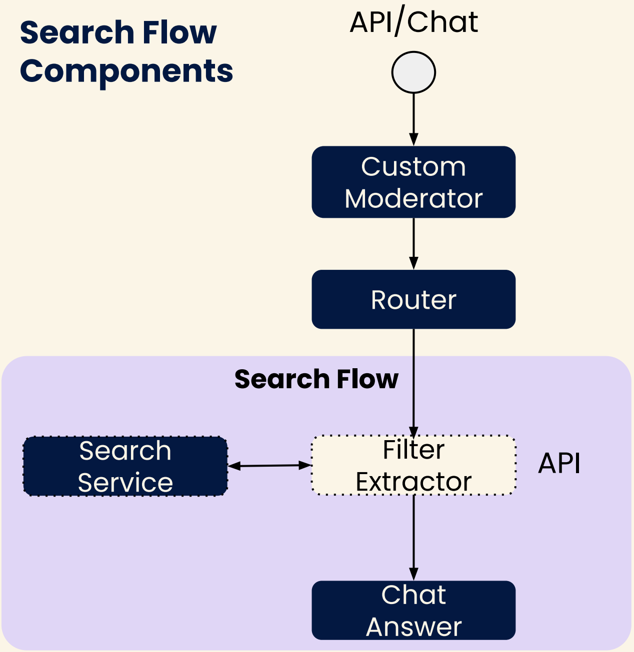


Chat. Solution



Chat. Solution

Search Flow Components



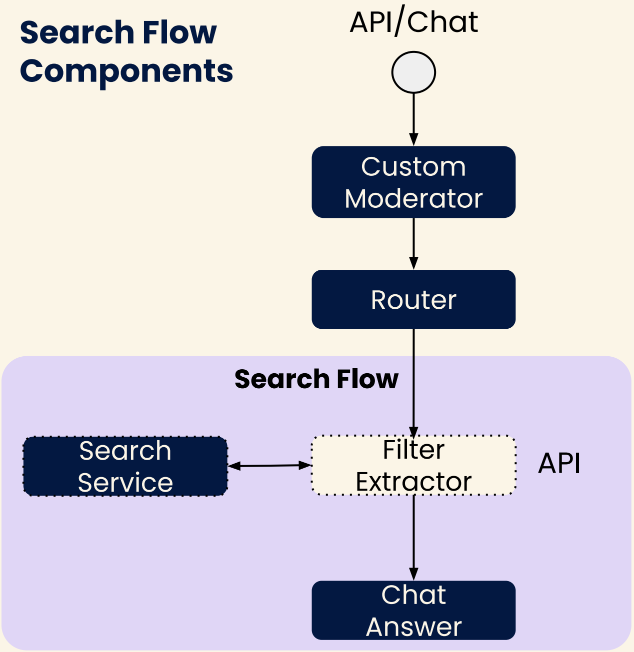
1 Filter Extractor

The heart of the system, that generate updated JSONs with user preferences based on user input. Is the interface with the Search Service.

2 Chat Answer

Tracks user preferences and extractor actions, communicates updates from the extractor, and guides the user in providing more data. Has some knowledge about locations.

Chat. Solution



1 Filter Extractor

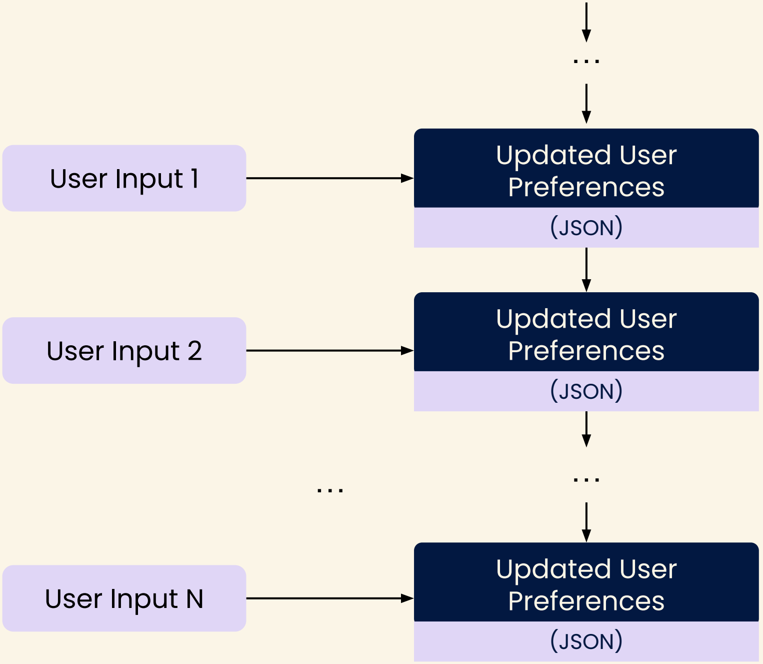
The heart of the system, that generate updated JSONs with user preferences based on user input. Is the interface with the Search Service.

2 Chat Answer

Tracks user preferences and extractor actions, communicates updates from the extractor, and guides the user in providing more data. Has some knowledge about locations.

Chat. Solution

Filter Extractor



Add New JSON Fields (Filters)

Updated JSON Fields (Filters)

Remove JSON Fields (Filters)

Chat. Solution

Filter Extractor

1 71 Different Filters (Classes)

Property type (House, Apartment, etc)

Business Context (Rent, Sale)

Amenities and Installations

Etc

2 Extract Visual Elements

Characteristics We Can Find in Images

Custom Moderation

Individual Extraction

Chat. Solution

Filter Extractor

1

71 Different Filters (Classes)

Property type (House, Apartment, etc)

Business Context (Rent, Sale)

Amenities and Installations

Etc

2

Extract Visual Elements

Characteristics We Can Find in Images

Custom Moderation

Individual Extraction

{
"**property_type**": "apartment",
"**max_price**" : 3000,
"**context**" :"rent",
"**bedrooms**": 3,
"**location**": "Sao Paulo, SP, Brazil",
"**visual_search**": ["modern decor"],
}

Chat. Solution

Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

Chat. Solution

Filter Extractor

User Current Preferences

{
  "**property_type**": "apartment",
  "**max_price**" : 3000,
  "**context**" :"rent",
  "**bedrooms**": 3,
  "**location**": "Sao Paulo, SP, Brazil",
  "**visual_search**": ["modern decor"],
}

User Textual Input

"Actually, 2 bedrooms and industrial style."

Chat. Solution

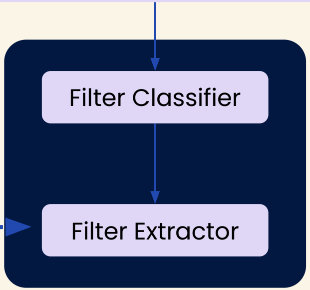
Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."



Two Steps
Dynamic
Extraction

Chat. Solution

Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."

Filter Classifier

Filter Extractor

Two Steps
Dynamic
Extraction

LLM Call: Detects Preferences Mentioned by the User
[bedrooms, visual search]

Chat. Solution

Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."

Filter Classifier

Filter Extractor

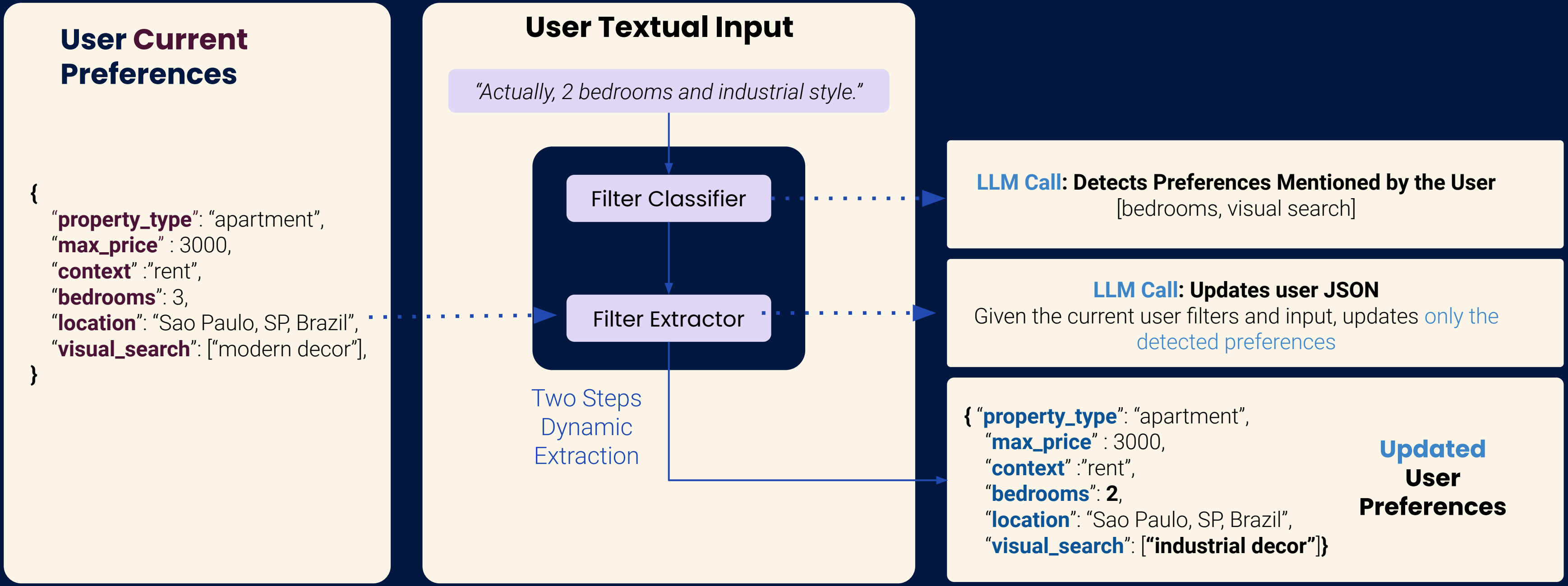
Two Steps Dynamic Extraction

LLM Call: Detects Preferences Mentioned by the User
[bedrooms, visual search]

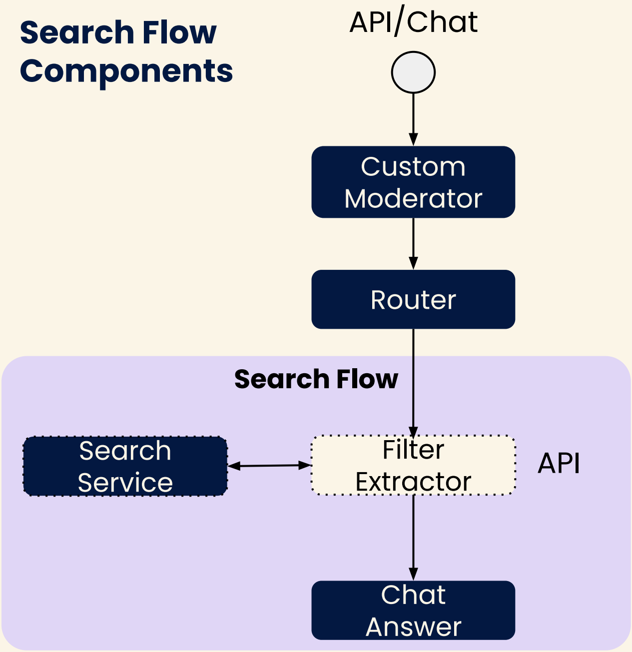
LLM Call: Updates user JSON
Given the current user filters and input, updates only the detected preferences

Chat. Solution

Filter Extractor



Chat. Solution



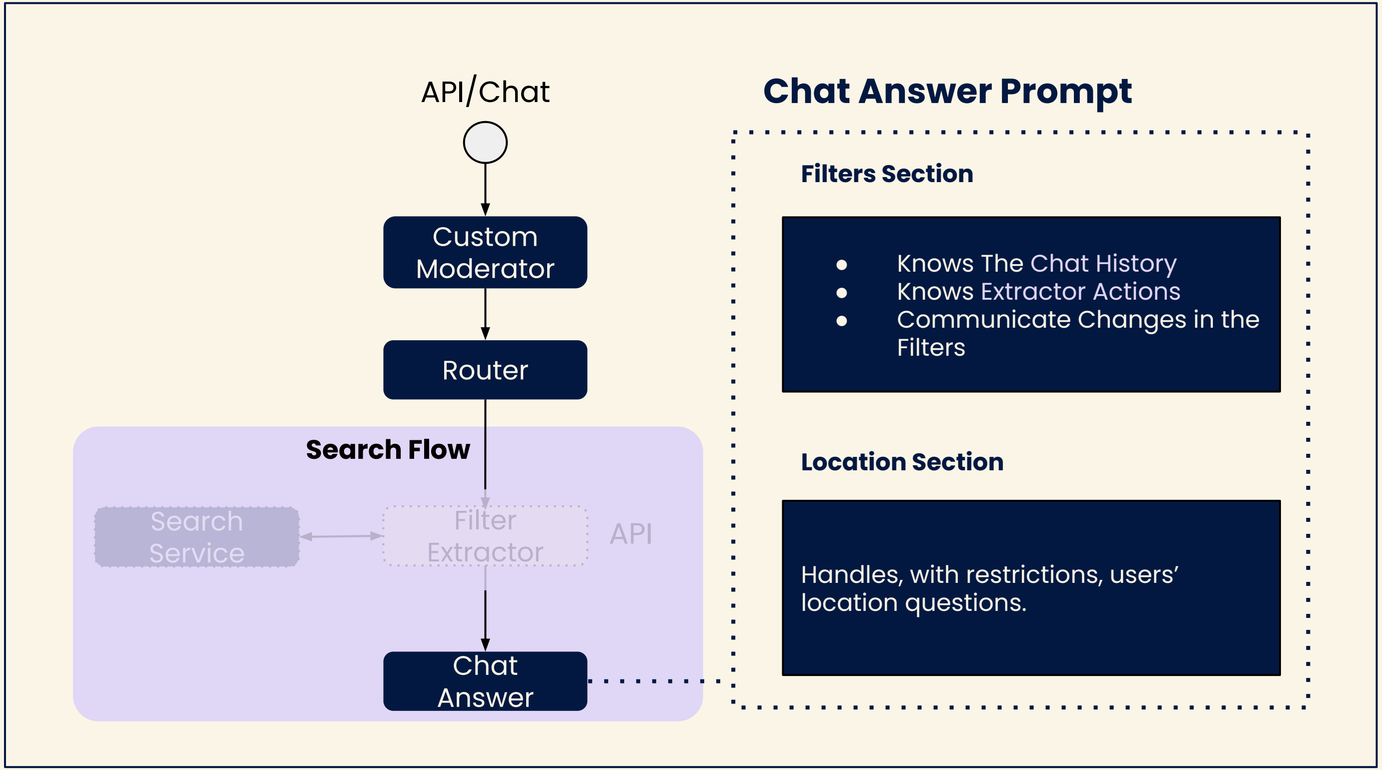
1 Filter Extractor

The heart of the system, that generate updated JSONs with user preferences based on user input. Is the interface with the Search Service.

2 Chat Answer

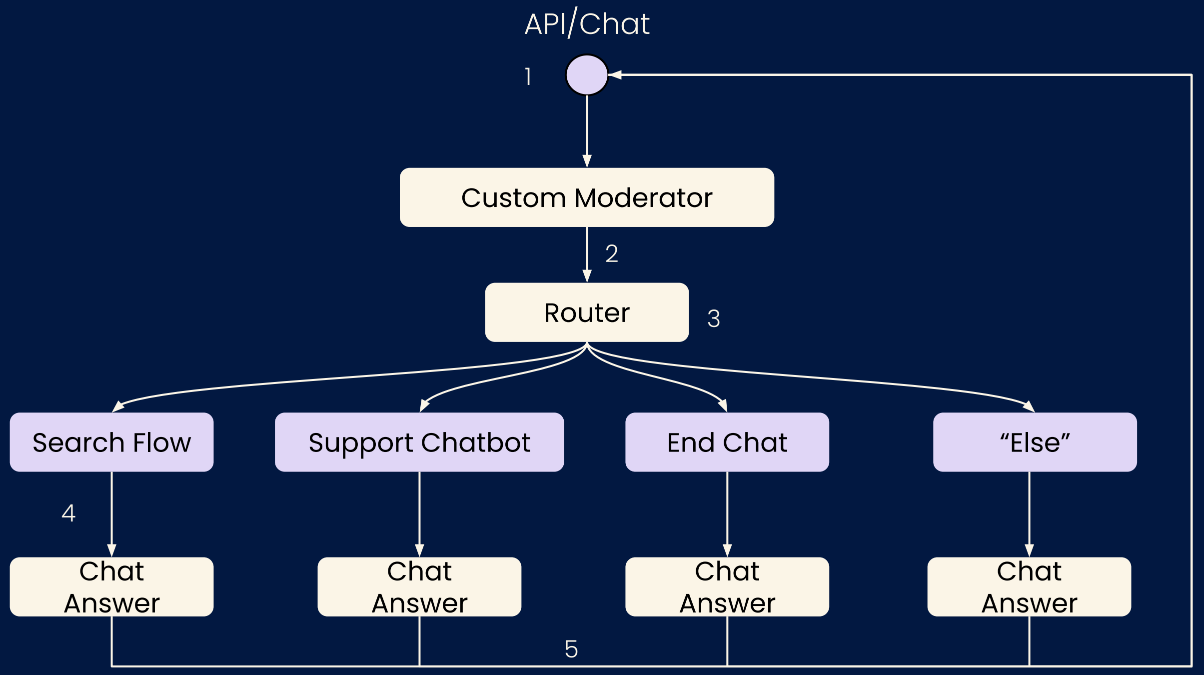
Tracks user preferences and extractor actions, communicates updates from the extractor, and guides the user in providing more data. Has some knowledge about locations.

Chat. Solution



But....is it good at all?

Chat. Evaluation



1. The user Input

2. If the input was moderated or not

3. The Router Predictions

4. Extracted Filters

5. Chat Answers

\+ LLM Adopted Prompts

Chat. Evaluation

Chat Component	Validation Type	Eval Dataset	Evaluator
Router	Offline/CICD	Golden Dataset	Classical M.L Metrics
	Online	Production Dataset	LLM As A Judge
Filter Extractor	Offline/CICD	Golden Dataset	Classical M.L Metrics
	Online	Production Dataset	LLM As A Judge
Visual Components	Offline/CICD	Golden Dataset	LLM As A judge
	Online	Production Dataset	
Chat Answers	Offline/CICD	Golden Dataset	LLM As A Judge
	Online	Production Dataset	

Chat. Evaluation

Classical ML Evaluation

- Objective/Deterministic Tasks, such as Classification (Filter Extractor, Router)
- Used Offline Only (CI/CD)
- More Granular/Detailed metrics (Filter-level)

LLM As a Judge Evaluation

- Subjective/Open-Ended Tasks, such as Conversation (Chat Answers)
- Used Online and Offline (CI/CD)
- General Metrics

Chat. Evaluation

Classical ML Evaluation

Example of Filter Extractor Metrics

Filter Group	Filter Category	F1	Precision	Recall
property_type	condominium	0.909	0.968	0.857
property_type	apartment	0.970	0.963	0.978
business_context	rent	0.991	0.988	0.994
business_context	sale	0.981	0.981	0.981

LLM As a Judge Evaluation

User Input: Santa Cecília 2 bedrooms

Score	1 (Correct)
Score Reason	The LLM correctly updated the location filter to 'Santa Cecilia, São Paulo - SP, Brasil' Already Existing filter for 2 bedrooms (min_bedrooms: 2), so no change was needed there.

Chat. Evaluation

LLM As A Judge

How We Do It

1	**Simplify The Task** Simplify the original task to make evaluation accurate
2	**Chain of Thoughts** A list of clear instructions
3	**Clear Grading Criteria** Clear grading guide to produce final score
3	**Score Reason** Ask the LLM to Produce a reason for the scoring

Chat. Evaluation

LLM As A Judge

How We Do It	
1	**Simplify The Task** Simplify the original task to make evaluation accurate
2	**Chain of Thoughts** A list of clear instructions
3	**Clear Grading Criteria** Clear grading guide to produce final score
3	**Score Reason** Ask the LLM to Produce a reason for the scoring

Example	
User Input	Santa Cecília 2 bedrooms
Input to Judge	**User Message** Santa Cecília 2 bedrooms **Current User Preferences** business_context: RENT location: Rua Doutor Rubens Meireles, São Paulo, SP, Brasil max_price: 5700.0 min_bedrooms: 2 **What Extractor Did** location CHANGED FROM *Rua Doutor Rubens Meireles, São Paulo, SP, Brasil* TO *Santa Cecilia, Sã Paulo - SP, Brasil.*

Chat. Evaluation

LLM As A Judge

How We Do It

1	**Simplify The Task** Simplify the original task to make evaluation accurate
2	**Chain of Thoughts** A list of clear instructions
3	**Clear Grading Criteria** Clear grading guide to produce final score
3	**Score Reason** Ask the LLM to Produce a reason for the scoring

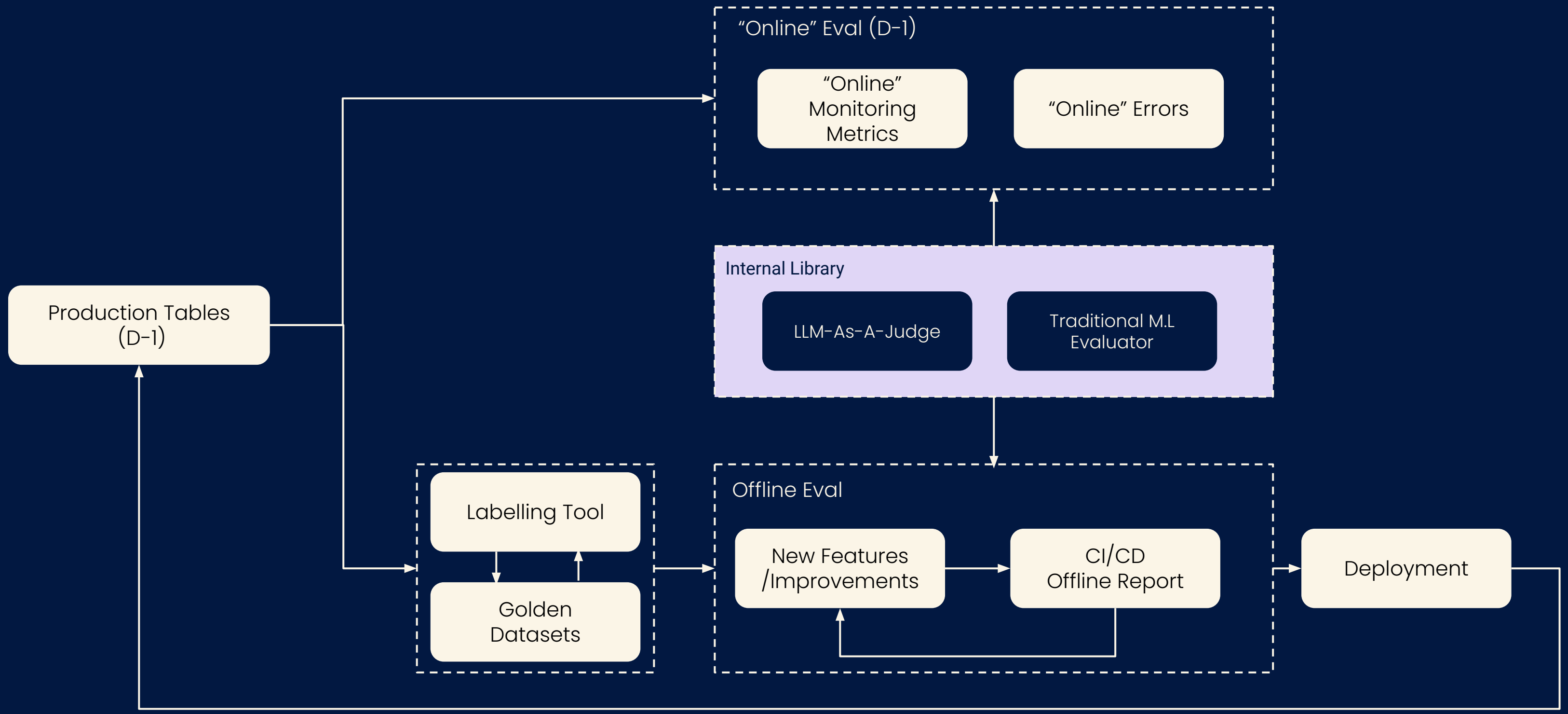
Example

User Input	Santa Cecília 2 bedrooms
Input to Judge	**User Message** Santa Cecília 2 bedrooms **Current User Preferences** business_context: RENT location: Rua Doutor Rubens Meireles, São Paulo, SP, Brasil max_price: 5700.0 min_bedrooms: 2 **What Extractor Did** location CHANGED FROM *Rua Doutor Rubens Meireles, São Paulo, SP, Brasil* TO *Santa Cecilia, Sã Paulo - SP, Brasil.*

Score	1 (Correct)
Score Reason	The LLM correctly updated the location filter to 'Santa Cecilia, São Paulo - SP, Brasil' Already Existing filter for 2 bedrooms (min_bedrooms: 2), so no change was needed there.

Chat. Evaluation

Overall Architecture



Chat. Solution

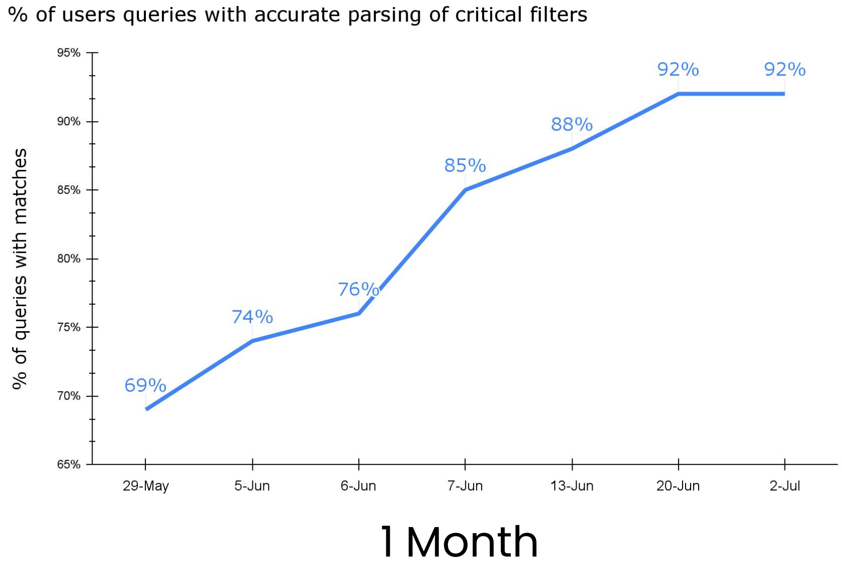
Lessons Learned

What We Tried		What Worked Best	
1	"One Prompt to Rule Them All"	1	Combination of LLM and State Machine
2	Granular Approach: - Break down text into topics; - Classify Each Topic; - Extraction for Each Topic;	2	Two Steps Approach and Individual Calls For Complex Filters: - Amenities and Installations - Visual Elements
3	Synthetic Datasets	3	Manually Generated Golden Datasets

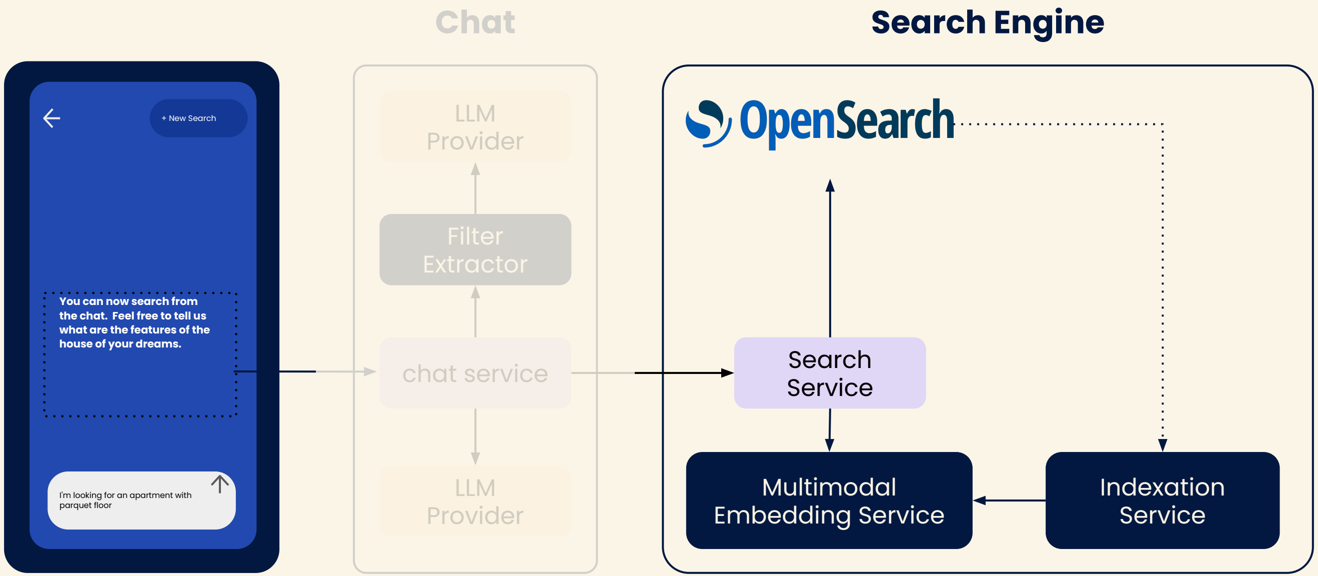
Chat. Solution

Lessons Learned

Right Evaluation Stimulates Virtuous Cycle of Improvements

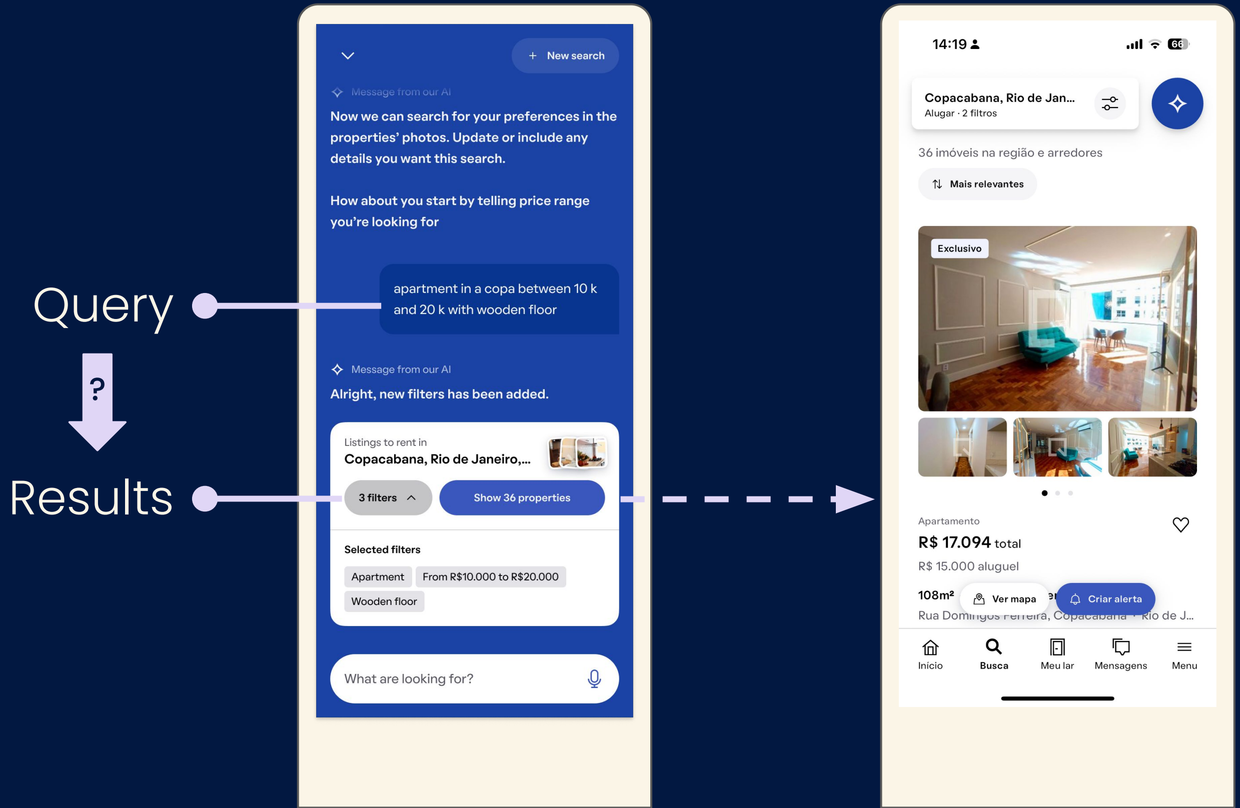


LLM-based Search Assistant



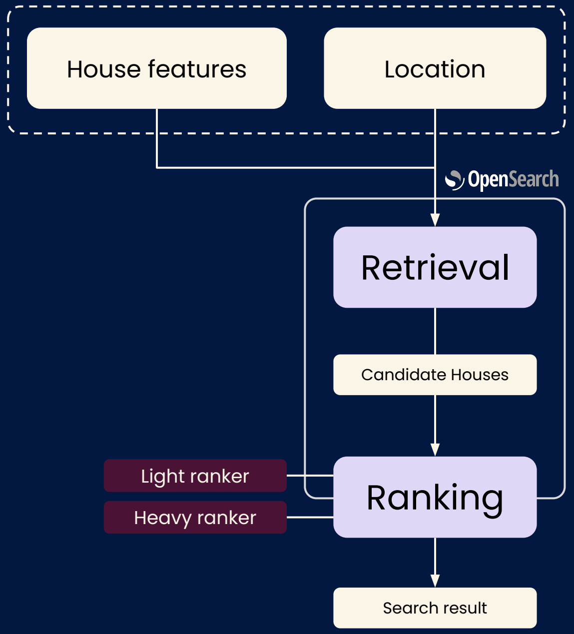
Search

Before the chat



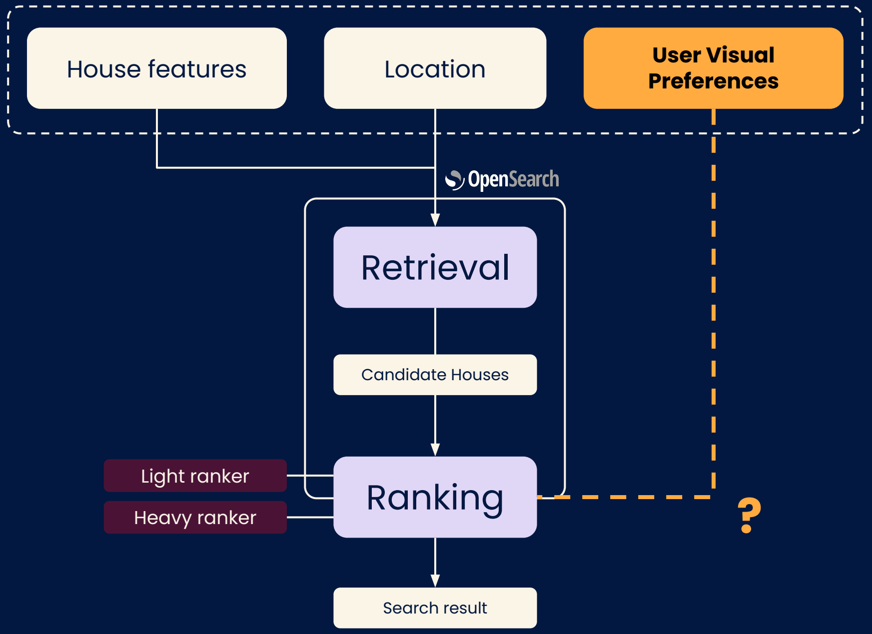
Search

Before the chat



Search

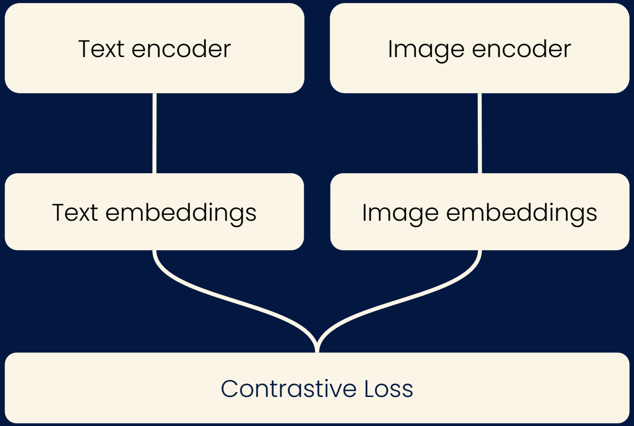
After the chat



Search

Multimodal embeddings

"wooden floor"



OpenAI CLIP

Learning Transferable Visual Models From Natural Language Supervision

Alec Radford * [1] **Jong Wook Kim** * [1] **Chris Hallacy** [1] **Aditya Ramesh** [1] **Gabriel Goh** [1] **Sandhini Agarwal** [1]
Girish Sastry [1] **Amanda Askell** [1] **Pamela Mishkin** [1] **Jack Clark** [1] **Gretchen Krueger** [1] **Ilya Sutskever** [1]

Google SigLIP

Sigmoid Loss for Language Image Pre-Training

Xiaohua Zhai* Basil Mustafa Alexander Kolesnikov Lucas Beyer*
Google DeepMind, Zürich, Switzerland
{xzhai, basilm, akolesnikov, lbeyer}@google.com

Salesforce BLIP-2

BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models

Junnan Li Dongxu Li Silvio Savarese Steven Hoi
Salesforce Research

Search

Scoring

User Query

*"I'm looking for a house with **wooden floor, big windows** and **modern decoration**"*

Preferences Extractor

Wooden Floor

Big Windows

Modern Decoration

Scoring Function

House A

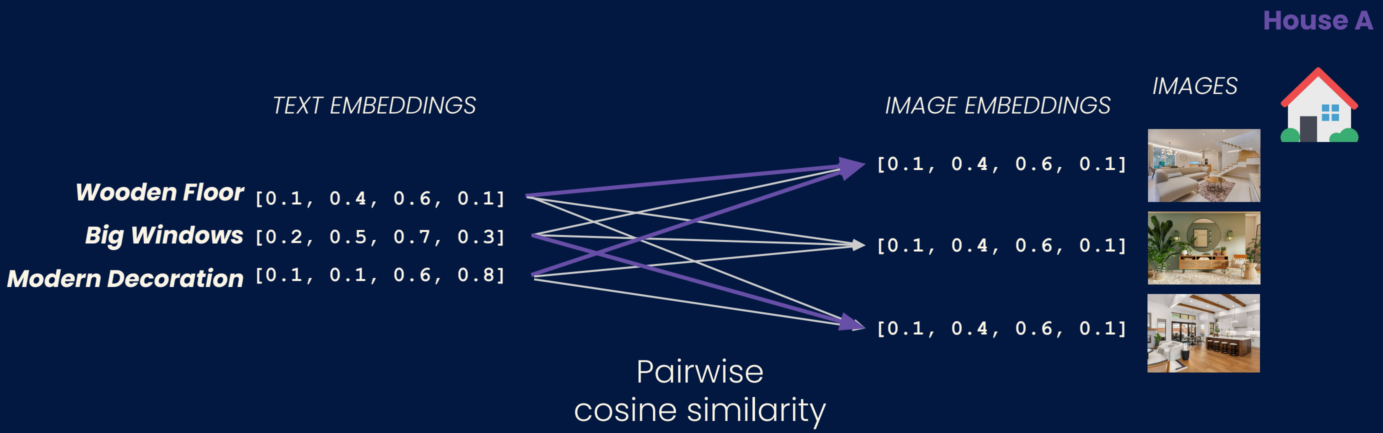
IMAGES

IMAGES

House B

Search

Scoring

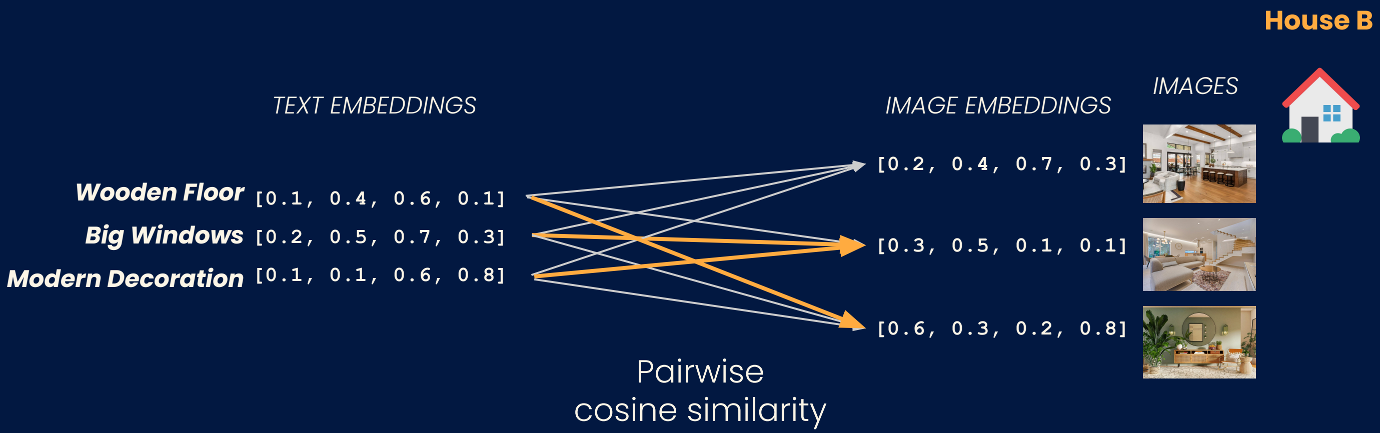


$$\text{IMAGE\_SCORE}_{\texttt{houseA}} = \textbf{sum}(\textbf{max}(\textbf{cosine})\ \textbf{group by}\ \textit{Visual Preference})$$

*For query "I'm looking for a house with **wooden floor, big windows** and **modern decoration**"*

Search

Scoring



$\text{IMAGE\_SCORE}_{\text{houseB}}$ = **sum**(**max**(**cosine**) **group by** *Visual Preference*)

For query "I'm looking for a house with ***wooden floor, big windows*** *and* ***modern decoration****"*

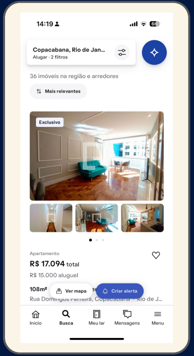
Search

Scoring

$$\mathrm{SCORE}_{\mathtt{houseA}} = \mathbf{combine}(\mathrm{IMAGE\_SCORE}_{\mathtt{houseA}}, \mathrm{HEAVY\_RANKER\_SCORE}_{\mathtt{houseA}}, \ldots)$$

$$\mathrm{SCORE}_{\mathtt{houseB}} = \mathbf{combine}(\mathrm{IMAGE\_SCORE}_{\mathtt{houseB}}, \mathrm{HEAVY\_RANKER\_SCORE}_{\mathtt{houseB}}, \ldots)$$

$$\mathrm{RESULTS} = \mathbf{topK}(\mathbf{rank}(\mathrm{SCORE}_{\mathtt{houseA}}, \mathrm{SCORE}_{\mathtt{houseB}}))$$



But....is it good at all?

Search

Evaluation

Traditional

search results

user interaction

+

Relevance metrics
precision@k, ndcg@k, hits@k, ...

Conversion metrics
search2visit, search2contract

Overall health of the system

LLM-as-a-judge

search results

user query

Did results address the user requirements?

Results (multimodal) explainability

Search

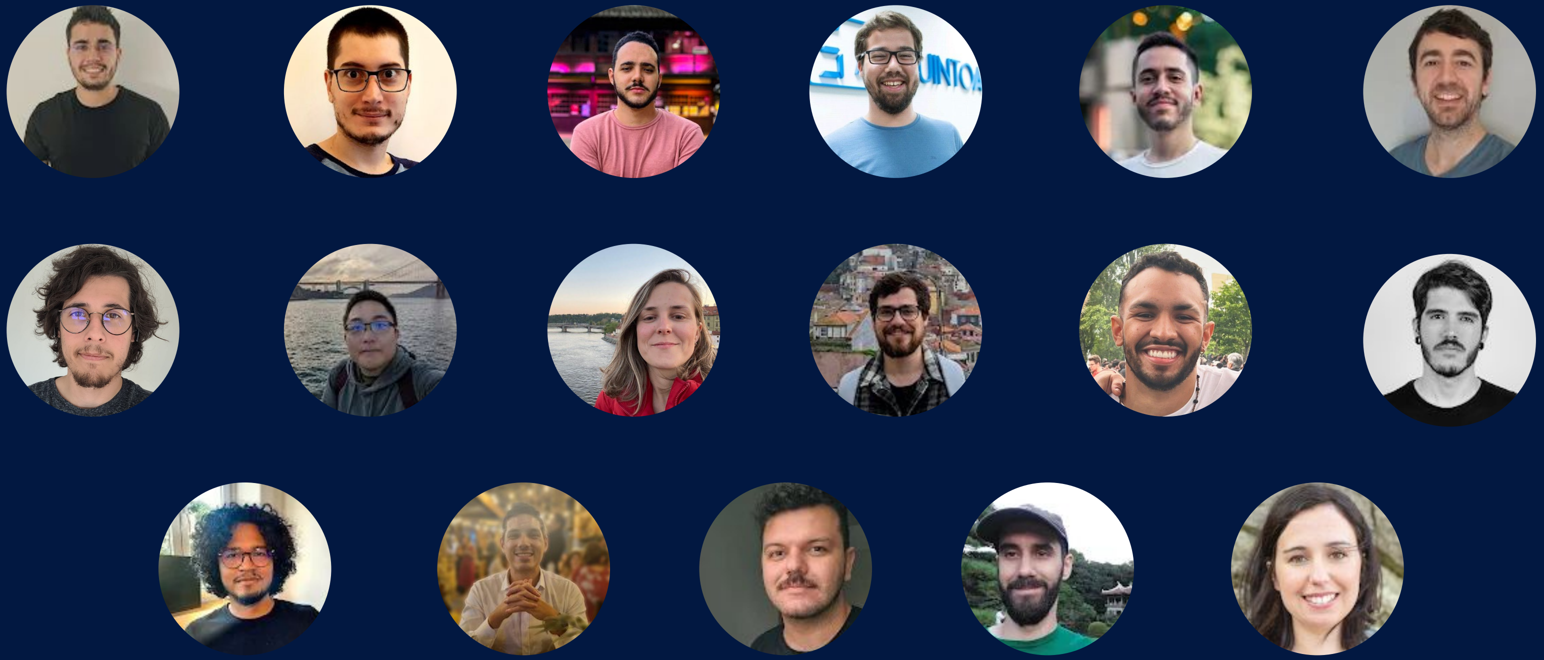
What we did

- Similarity-based image sampling
- Different similarity aggregation strategies
  - AVG(SUM(cosine))
  - AVG(MAX(cosine))
  - MAX(MAX(cosine))

Future and WIP

- Hybrid Search: full-text + vector search
- House-level vs. image-level embeddings
- Heavy ranker with image features
- Visual features as filter to reduce the search space

Thank You Team



We are hiring

Come join our team!

https://carreiras.quintoandar.com.br/?lang=en


Grupo
QuintoAndar

Agenda

1. Who We Are
2. Why a Search Assistant?
3. What it is
4. The Chat Component
5. The Search Component
6. Final Remarks

Who are we?

We are QuintoAndar Group, the **largest real estate ecosystem in Latin America.**

Driven by a shared purpose of helping people love where they live, we offer a diverse portfolio of brands and solutions that cover all stages of the living journey.

We develop technologies and innovations to transform and enhance the housing experience as a whole.

Where are we?

Headquartered in Brazil and with a presence in five other Latin American countries (in addition to a tech hub in Europe), we are a global group that transcends borders.

Grupo QuintoAndar

The QuintoAndar Group currently operates in the following countries: **Brazil, Argentina, Ecuador, Panama, Peru, and Mexico.**

Our team is made up of talented professionals who work from various locations around the world.

Our Numbers

We have formed the largest real estate group in Latin America, and the numbers we've achieved help to illustrate the scale of our brand.

Grupo QuintoAndar

Traffic

45 million monthly visits to our platforms

Listings

Over 8 million property listings published

Assets

$30 billion USD in assets under management

Administration

Over 270,000 contracts under administration

What is the Product

Chat

Real estate agent would talk to you until they **understand what you need...**

So does Copilot!

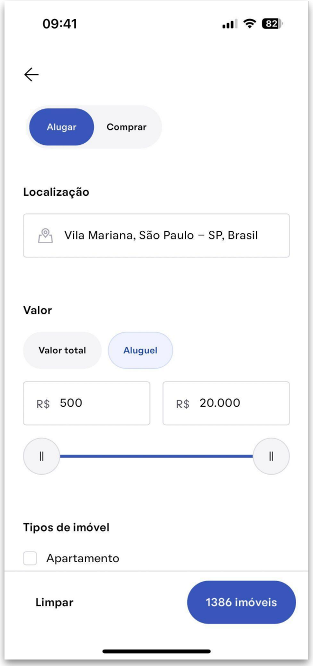
Search

Once a real estate agent understands your need, they will **find** the property **tailored** to your needs, prioritizing things most important for **you...**

Similar as Copilot!

Why did we need LLM-based Search Assistant? BEFORE

Search queries are filter-based only



Does apartment have wooden floor?

Need to review all photos manually

Is apartment close to supermarkets?

Need to go to the map and check

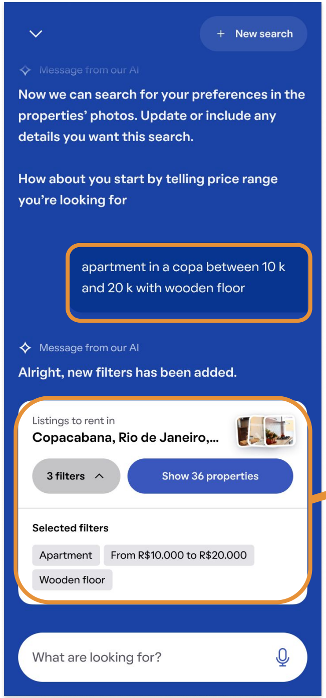
Me and my husband have 2 kids, we need to find nice apartment to fit us all close to the good schools

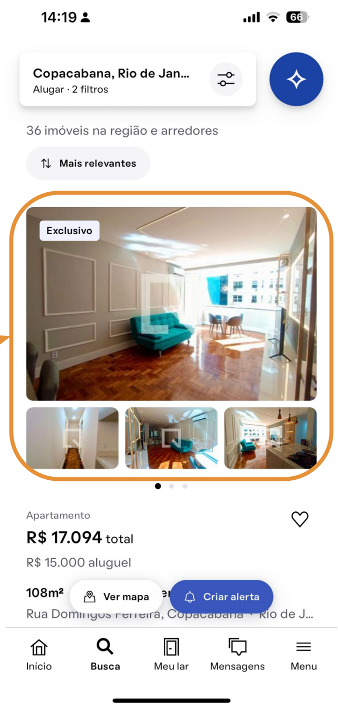
Need to adapt requirements to the filter structure

Why did we need LLM-based Search Assistant? NOW

Chat

Search





LLM-based Search Assistant

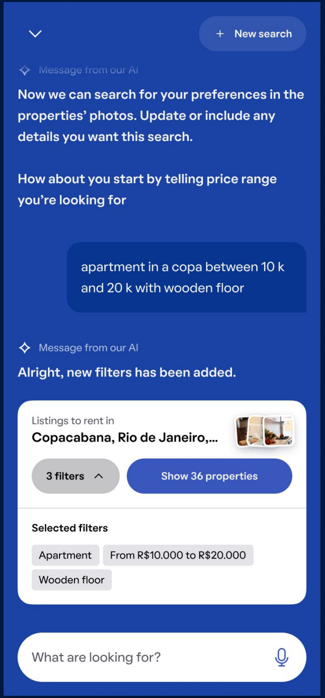
What it is

1 A Data Collector Chat

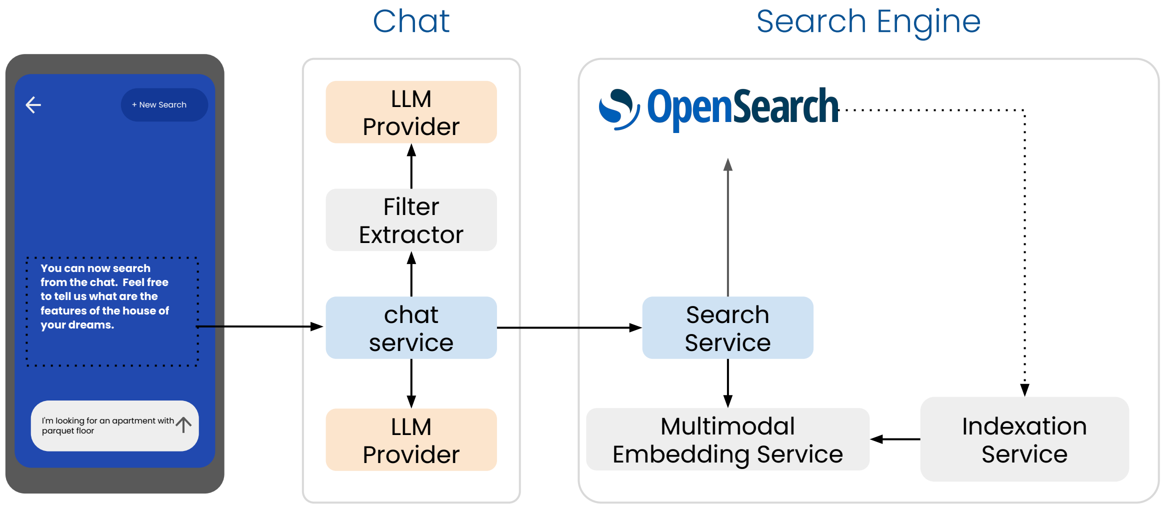
A chat that prompts the user to provide preferences until enough data is gathered to perform a search

2 A Search Engine

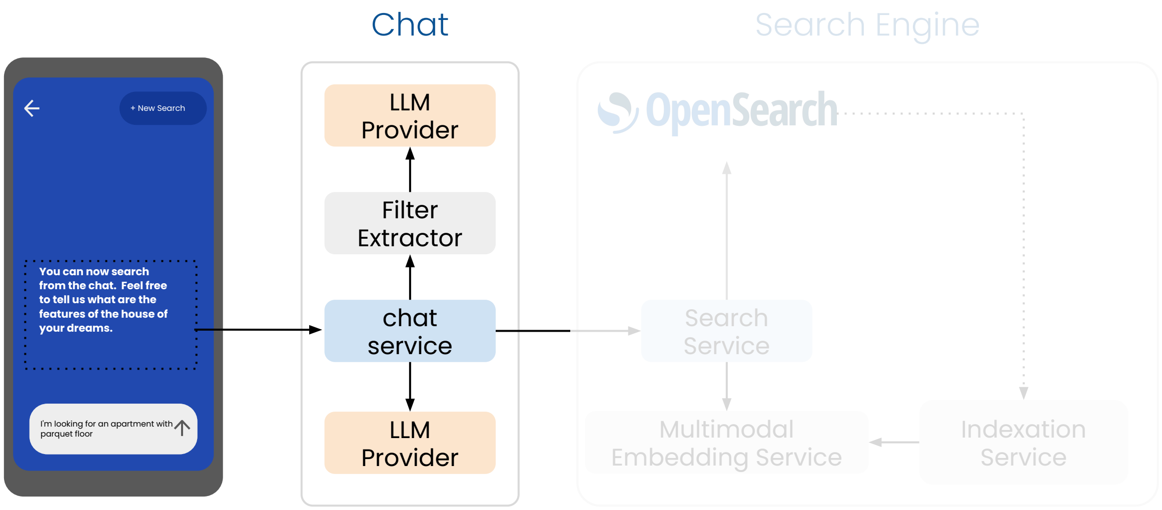
A tool that converts the user's preferences, including visual preferences, into a search.



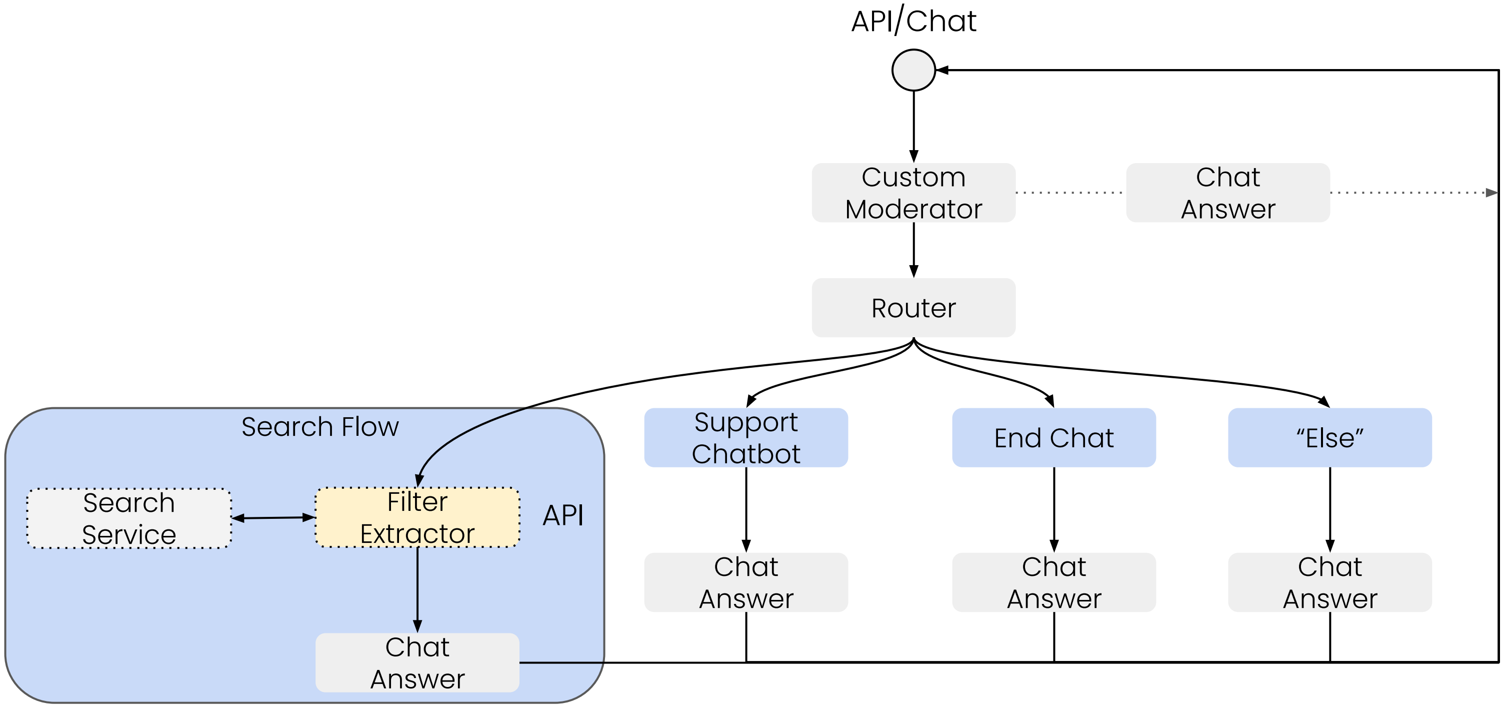
LLM-based Search Assistant



LLM-based Search Assistant

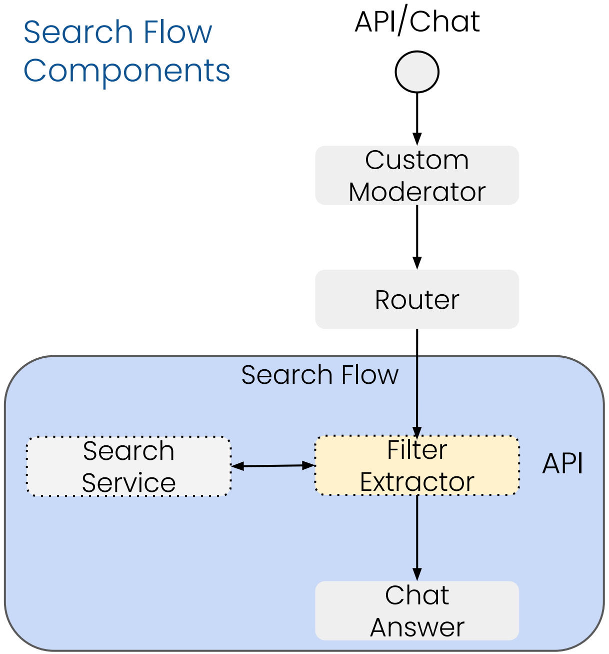


Chat. Solution



Chat. Solution

Search Flow Components



1 Filter Extractor

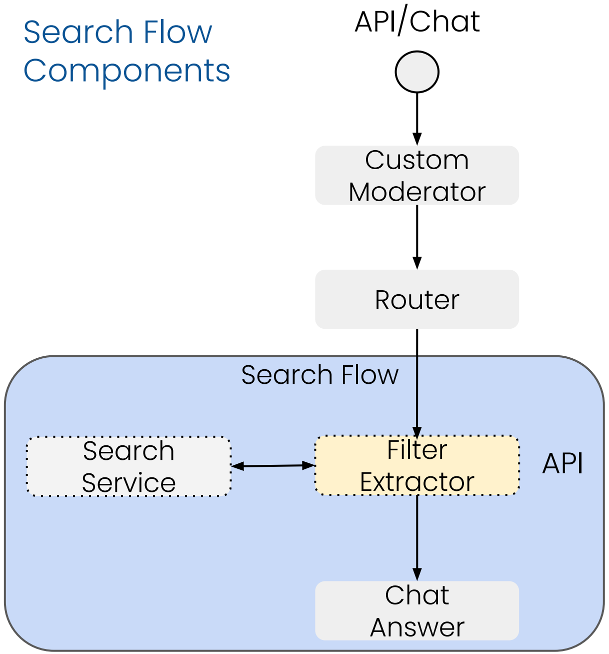
The **heart** of the system, that generate **updated** JSONs with user preferences based on user input. Is the interface with the Search Service.

2 Chat Answer

Tracks user preferences and extractor actions, communicates updates from the extractor, and guides the user in providing more data. Has some knowledge about locations.

Chat. Solution

Search Flow Components



1 Filter Extractor

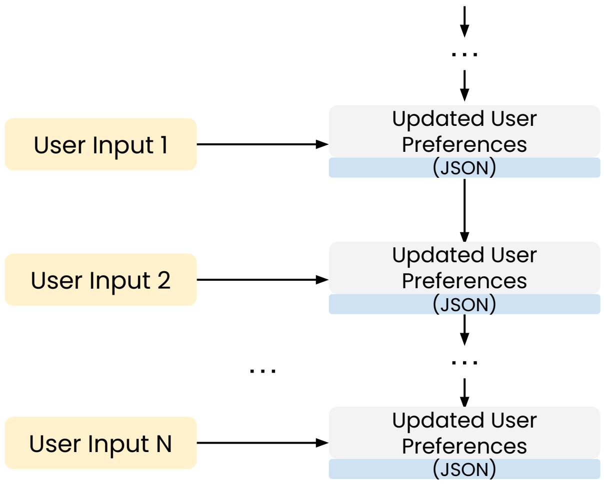
The **heart** of the system, that generate **updated** JSONs with user preferences based on user input. Is the interface with the Search Service.

2 Chat Answer

Tracks user preferences and extractor actions, communicates updates from the extractor, and guides the user in providing more data. Has some knowledge about locations.

Chat. Solution

Filter Extractor



Add New JSON Fields (Filters)

Updated JSON Fields (Filters)

Remove JSON Fields (Filters)

Chat. Solution

Filter Extractor

1 **71 Different Filters (Classes)**

- Property type (House, Apartment, etc)
- Business Context (Rent, Sale)
- Amenities and Installations
- Etc

2 **Extract Visual Elements**

- Characteristics We Can Find in Images
- Custom Moderation
- Individual Extraction

Chat. Solution

Filter Extractor

1 **71 Different Filters (Classes)**

Property type (House, Apartment, etc)

Business Context (Rent, Sale)

Amenities and Installations

Etc

2 **Extract Visual Elements**

Characteristics We Can Find in Images

Custom Moderation

Individual Extraction

{
“**property_type**”: “apartment”,
“**max_price**” : 3000,
“**context**” :”rent”,
“**bedrooms**”: 3,
“**location**”: “Sao Paulo, SP, Brazil”,
“**visual_search**”: [“modern decor”],
}

Chat. Solution

Filter Extractor

User Current Preferences

{
  "**property_type**": "apartment",
  "**max_price**" : 3000,
  "**context**" :"rent",
  "**bedrooms**": 3,
  "**location**": "Sao Paulo, SP, Brazil",
  "**visual_search**": ["modern decor"],
}

Chat. Solution

Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."

Chat. Solution

Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."

Filter Classifier

Filter Extractor

Two Steps
Dynamic
Extraction

Chat. Solution
Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."

Filter Classifier

Filter Extractor

Two Steps Dynamic Extraction

LLM Call: Detects Preferences Mentioned by the User
[bedrooms, visual search]

Chat. Solution

Filter Extractor

User Current Preferences

```
{
  "property_type": "apartment",
  "max_price" : 3000,
  "context" :"rent",
  "bedrooms": 3,
  "location": "Sao Paulo, SP, Brazil",
  "visual_search": ["modern decor"],
}
```

User Textual Input

"Actually, 2 bedrooms and industrial style."

Filter Classifier

Filter Extractor

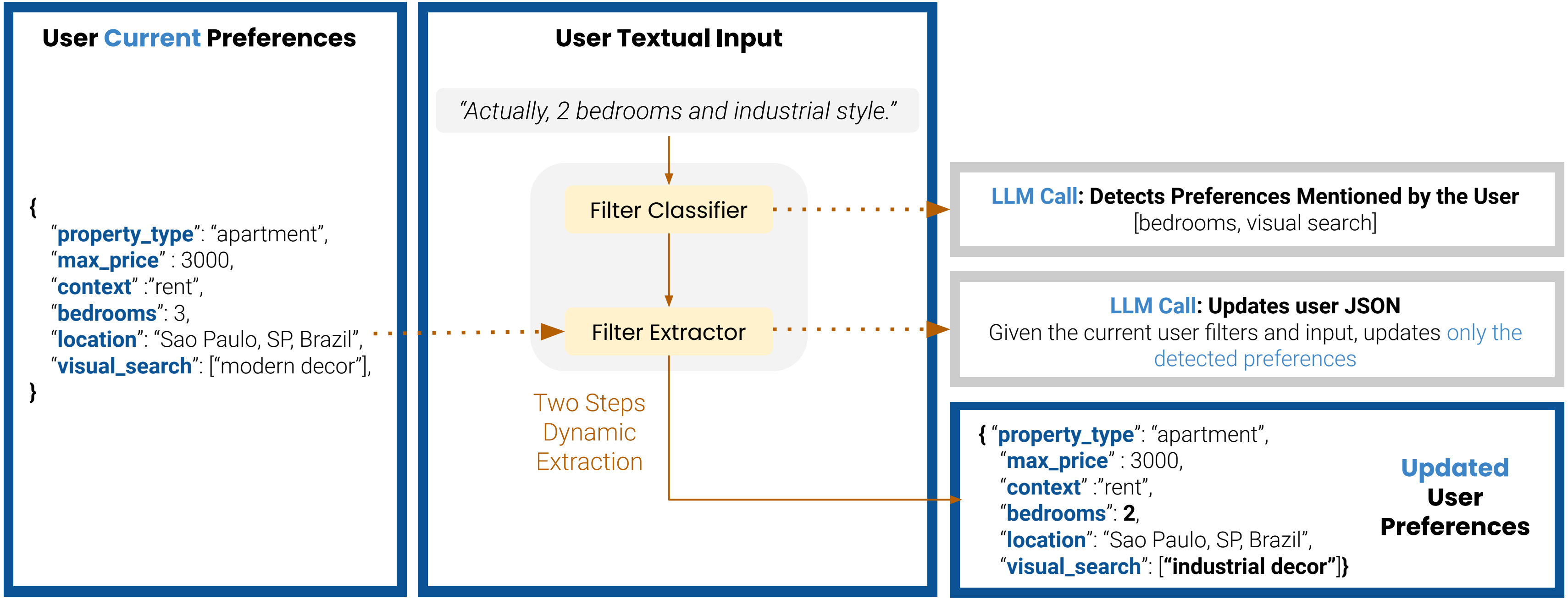
Two Steps Dynamic Extraction

LLM Call: Detects Preferences Mentioned by the User
[bedrooms, visual search]

LLM Call: Updates user JSON
Given the current user filters and input, updates only the detected preferences

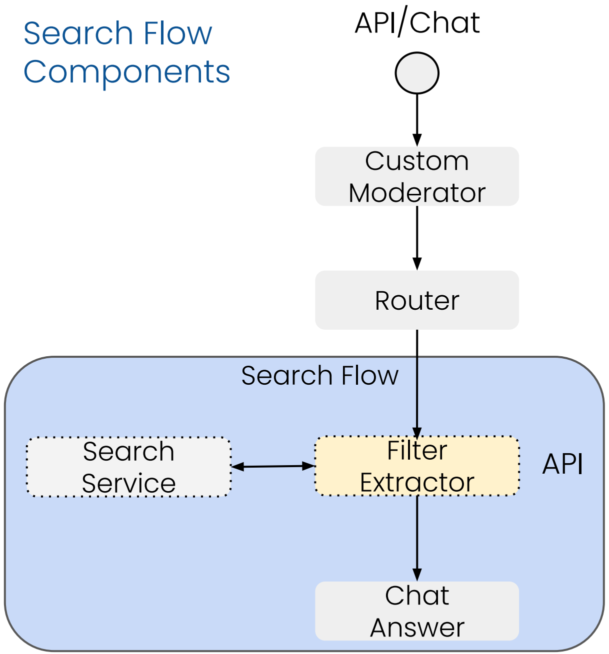
Chat. Solution

Filter Extractor



Chat. Solution

Search Flow Components



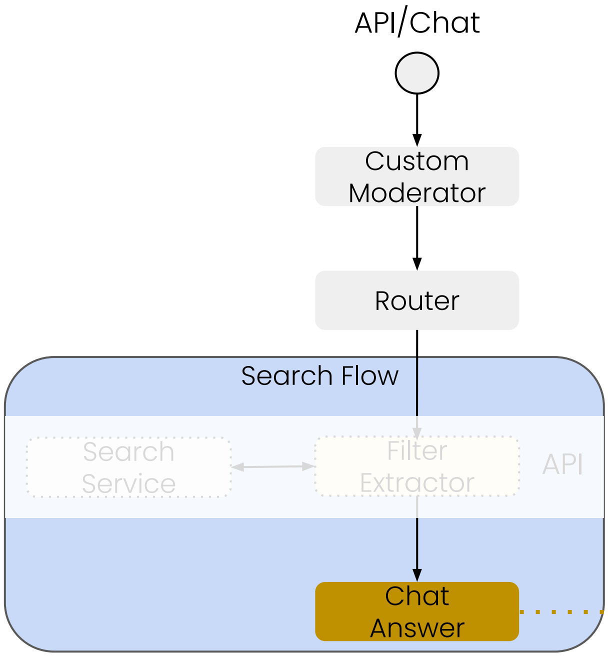
1 Filter Extractor

The **heart** of the system, that generate **updated** JSONs with user preferences based on user input. Is the interface with the Search Service.

2 Chat Answer

Tracks user preferences and extractor actions, communicates updates from the extractor, and guides the user in providing more data. Has some knowledge about locations.

Chat. Solution



Chat Answer Prompt

Filters Section

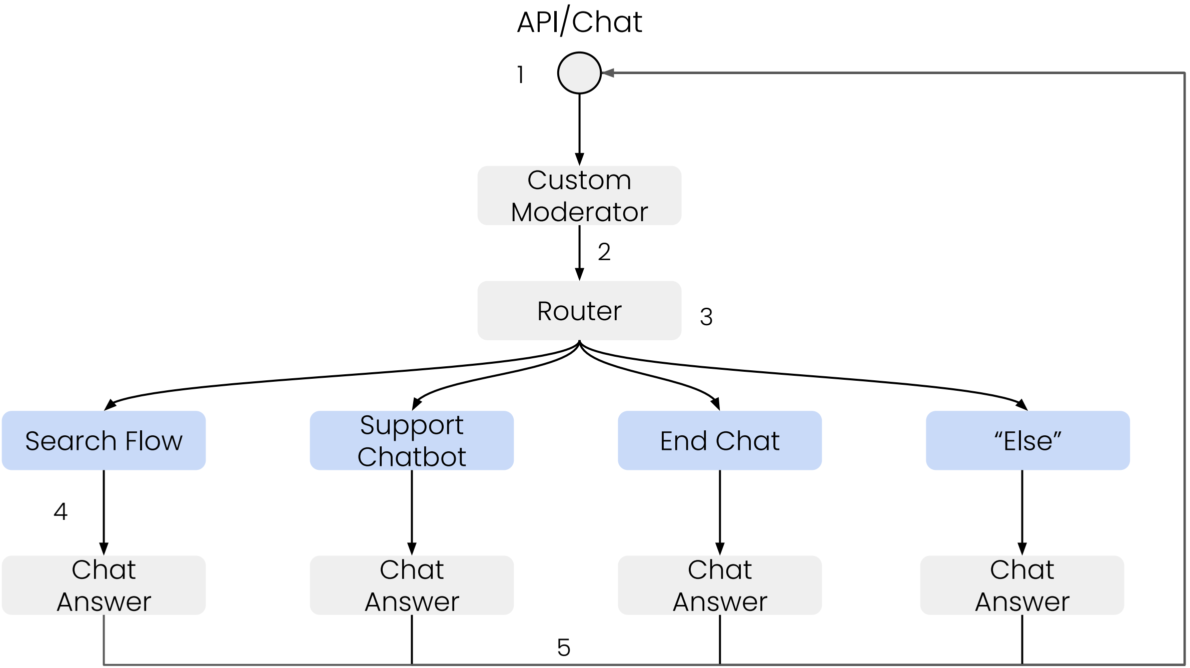
- Knows The Chat History
- Knows Extractor Actions
- Communicate Changes in the Filters

Location Section

Handles, with restrictions, users' location questions.

But....is it good at all?

Chat. Evaluation



1: The user Input

2: If the input was moderated or not

3: The Router Predictions

4: Extracted Filters

5: Chat Answers

+ LLM Adopted Prompts

Chat. Evaluation

Chat Component	Validation Type	Eval Dataset	Evaluator
Router	Offline/CICD	Golden Dataset	Classical M.L Metrics
	Online	Production Dataset	LLM As A Judge
Filter Extractor	Offline/CICD	Golden Dataset	Classical M.L Metrics
	Online	Production Dataset	LLM As A Judge
Visual Components	Offline/CICD	Golden Dataset	LLM As A judge
	Online	Production Dataset	
Chat Answers	Offline/CICD	Golden Dataset	LLM As A Judge
	Online	Production Dataset	

Chat. Evaluation

Classical ML Evaluation

- Objective/Deterministic Tasks, such as Classification (Filter Extractor, Router)
- Used Offline Only (CI/CD)
- More Granular/Detailed metrics (Filter-level)

LLM As a Judge Evaluation

- Subjective/Open-Ended Tasks, such as Conversation (Chat Answers)
- Used Online and Offline (CI/CD)
- General Metrics

Chat. Evaluation

Classical ML Evaluation

Example of Filter Extractor Metrics

Filter Group	Filter Category	F1	Precision	Recall
property_type	condominium	0.909	0.968	0.857
property_type	apartment	0.970	0.963	0.978
business_context	rent	0.991	0.988	0.994
business_context	sale	0.981	0.981	0.981

LLM As a Judge Evaluation

User Input: Santa Cecília 2 bedrooms

Score	1 (Correct)
Score Reason	- The LLM correctly updated the location filter to 'Santa Cecilia, São Paulo - SP, Brasil' - Already Existing filter for 2 bedrooms (min_bedrooms: 2), so no change was needed there.

Chat. Evaluation

LLM As A Judge

How We Do It

1 **Simplify The Task**
Simplify the original task to make evaluation accurate

2 **Chain of Thoughts**
A list of clear instructions

3 **Clear Grading Criteria**
Clear grading guide to produce final score

4 **Score Reason**
Ask the LLM to Produce a reason for the scoring

Chat. Evaluation

LLM As A Judge

How We Do It

1	**Simplify The Task** Simplify the original task to make evaluation accurate
2	**Chain of Thoughts** A list of clear instructions
3	**Clear Grading Criteria** Clear grading guide to produce final score
4	**Score Reason** Ask the LLM to Produce a reason for the scoring

Example

User Input	Santa Cecília 2 bedrooms
Input to Judge	**User Message** Santa Cecília 2 bedrooms **Current User Preferences** **business_context**: RENT **location**: Rua Doutor Rubens Meireles, São Paulo, SP, Brasil **max_price**: 5700.0 **min_bedrooms**: 2 **What Extractor Did** **location** CHANGED FROM *Rua Doutor Rubens Meireles, São Paulo, SP, Brasil* TO *Santa Cecilia, Sã Paulo - SP, Brasil*.

Chat. Evaluation

LLM As A Judge

How We Do It

1	**Simplify The Task** Simplify the original task to make evaluation accurate
2	**Chain of Thoughts** A list of clear instructions
3	**Clear Grading Criteria** Clear grading guide to produce final score
4	**Score Reason** Ask the LLM to Produce a reason for the scoring

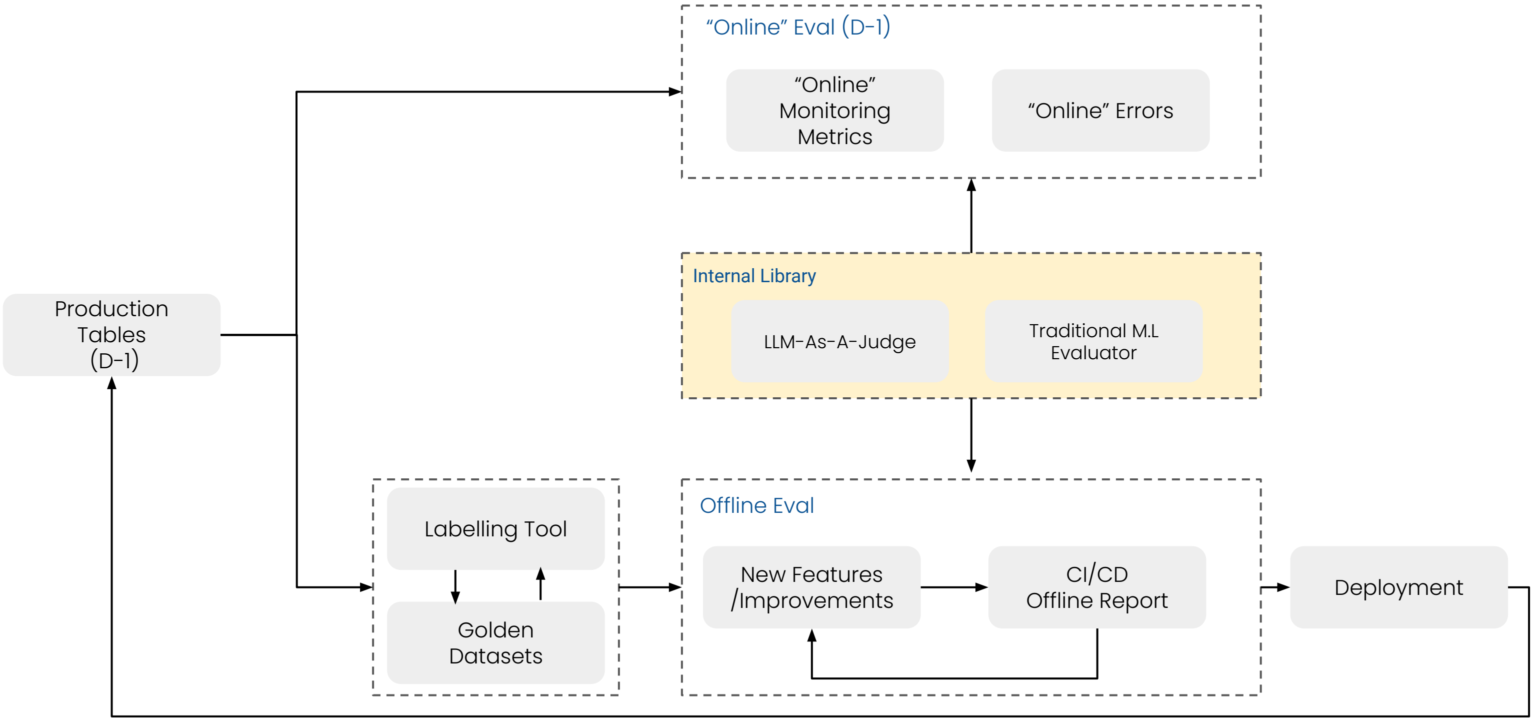
Example

User Input	Santa Cecília 2 bedrooms
Input to Judge	**User Message** Santa Cecília 2 bedrooms **Current User Preferences** **business_context**: RENT **location**: Rua Doutor Rubens Meireles, São Paulo, SP, Brasil **max_price**: 5700.0 **min_bedrooms**: 2 **What Extractor Did** **location** CHANGED FROM *Rua Doutor Rubens Meireles, São Paulo, SP, Brasil* TO *Santa Cecilia, Sã Paulo - SP, Brasil***.**

Score	1 (Correct)
Score Reason	The LLM correctly updated the location filter to 'Santa Cecilia, São Paulo - SP, Brasil' Already Existing filter for 2 bedrooms (min_bedrooms: 2), so no change was needed there.

Chat. Evaluation

Overall Architecture



Chat. Solution

Lessons Learned

What We Tried

1	"One Prompt to Rule Them All"
2	Granular Approach: – Break down text into topics; – Classify Each Topic; – Extraction for Each Topic;
3	Synthetic Datasets

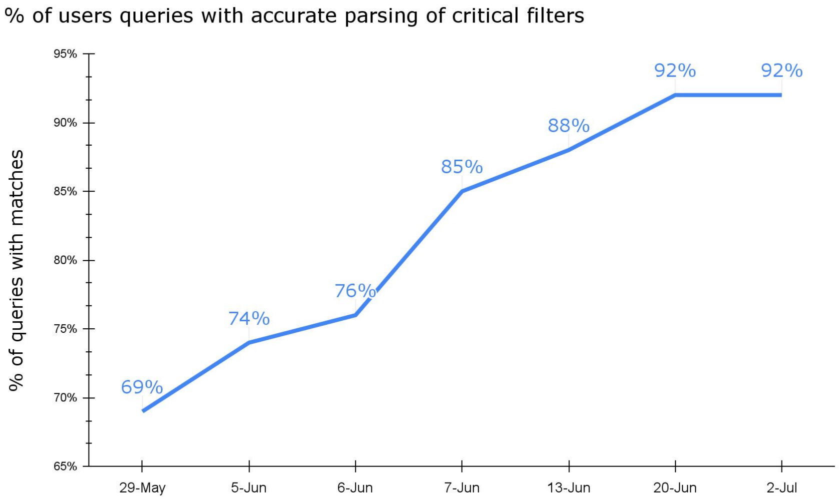
What Worked Best

1	Combination of LLM and State Machine
2	Two Steps Approach and Individual Calls For Complex Filters: – Amenities and Installations – Visual Elements
3	Manually Generated Golden Datasets

Chat. Solution

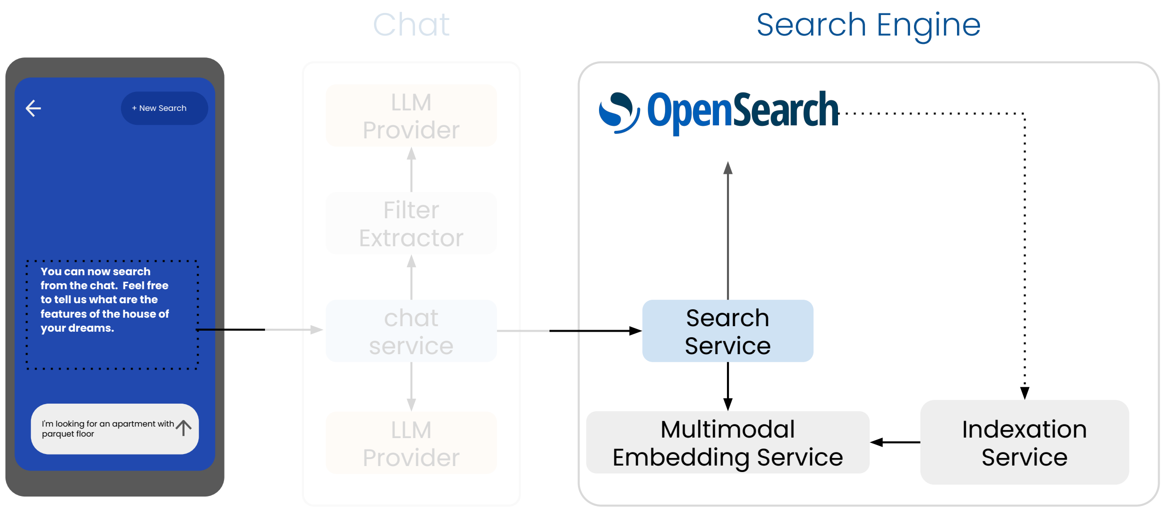
Lessons Learned

Right Evaluation Stimulates Virtuous Cycle of Improvements

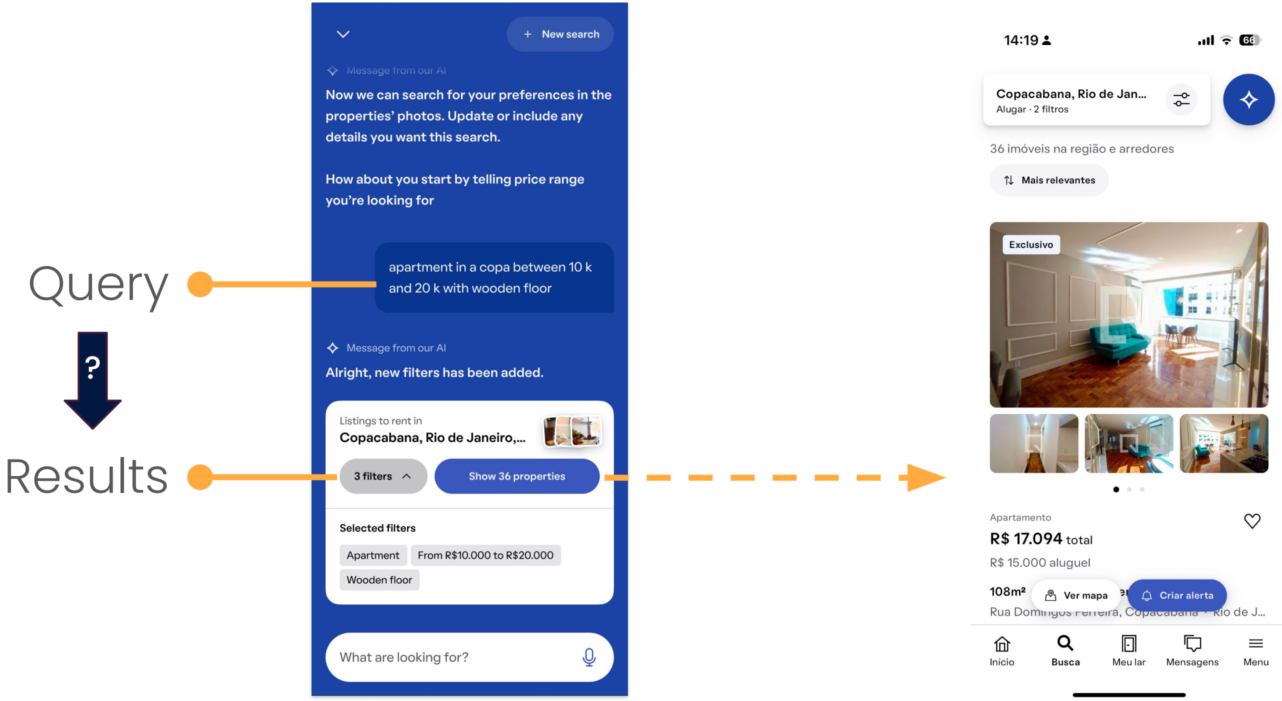


1 Month

LLM-based Search Assistant



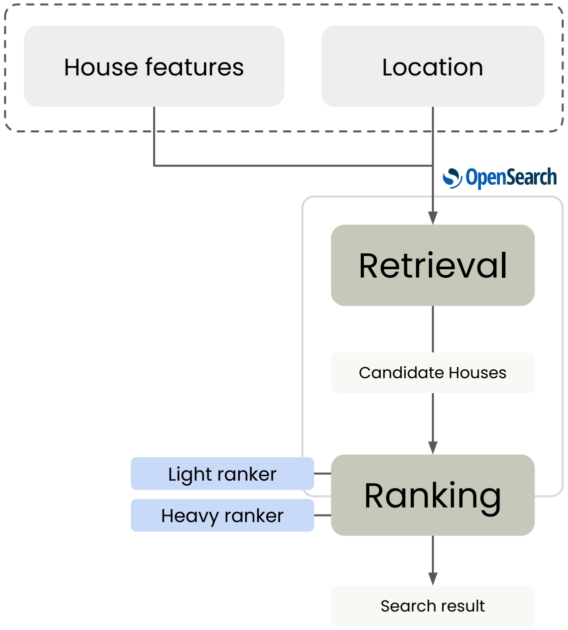
Search



Query
?
Results
New search
Message from our AI
Now we can search for your preferences in the properties' photos. Update or include any details you want this search.
How about you start by telling price range you're looking for
apartment in a copa between 10 k and 20 k with wooden floor
Message from our AI
Alright, new filters has been added.
Listings to rent in
Copacabana, Rio de Janeiro,...
3 filters
Show 36 properties
Selected filters
Apartment
From R$10.000 to R$20.000
Wooden floor
What are looking for?
14:19
Copacabana, Rio de Jan...
Alugar · 2 filtros
36 imóveis na região e arredores
Mais relevantes
Exclusivo
Apartamento
R$ 17.094 total
R$ 15.000 aluguel
108m²
Ver mapa
Criar alerta
Início
Busca
Meu lar
Mensagens
Menu

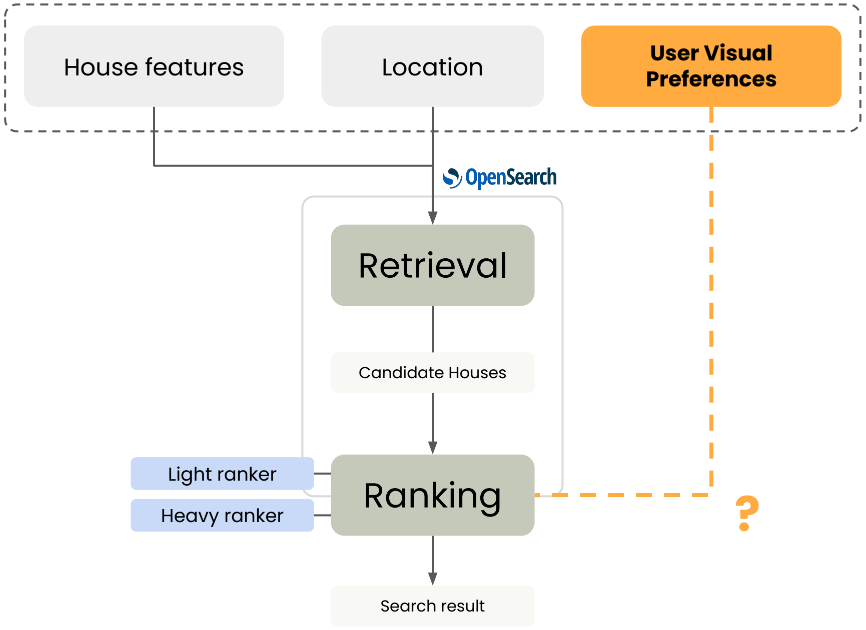
Search

Before the chat



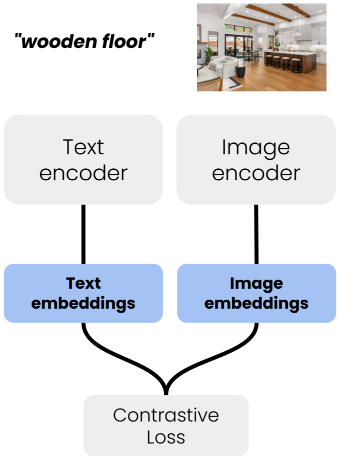
Search

After the chat



Search

Multimodal embeddings



OpenAI CLIP

Learning Transferable Visual Models From Natural Language Supervision

Alec Radford [*1] Jong Wook Kim [*1] Chris Hallacy [1] Aditya Ramesh [1] Gabriel Goh [1] Sandhini Agarwal [1] Girish Sastry [1] Amanda Askell [1] Pamela Mishkin [1] Jack Clark [1] Gretchen Krueger [1] Ilya Sutskever [1]

Google SigLIP

Sigmoid Loss for Language Image Pre-Training

Xiaohua Zhai* Basil Mustafa Alexander Kolesnikov Lucas Beyer*
Google DeepMind, Zürich, Switzerland
{xzhai, basilm, akolesnikov, lbeyer}@google.com

Salesforce BLIP-2

BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models

Junnan Li Dongxu Li Silvio Savarese Steven Hoi
Salesforce Research

Search

Scoring

User Query

"I'm looking for a house with ***wooden floor, big windows*** *and* ***modern decoration****"*

Preferences Extractor

Wooden Floor

Big Windows

Modern Decoration

Scoring Function

House A

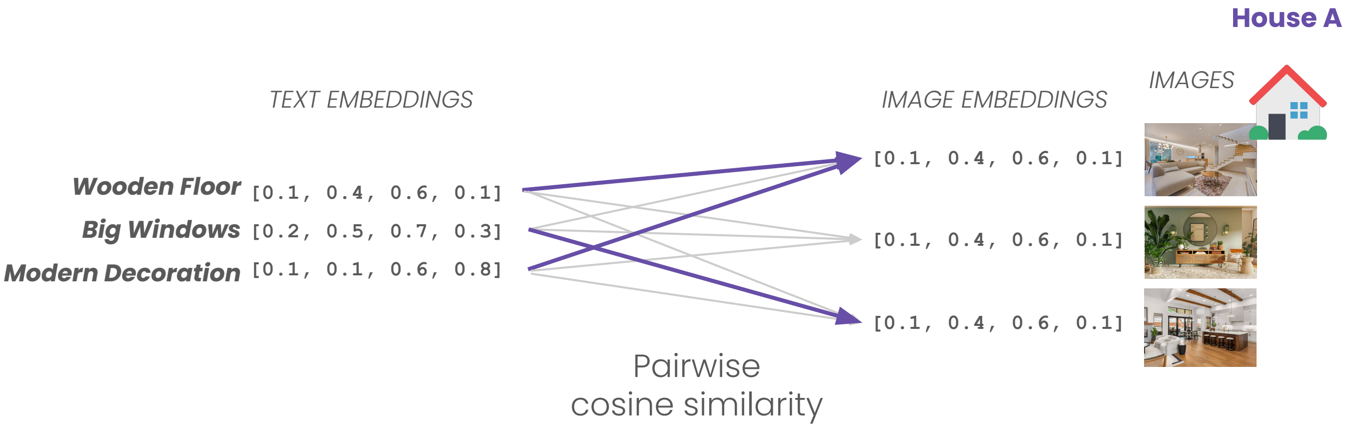
IMAGES

IMAGES

House B

Search

Scoring



IMAGE_SCORE$_{\texttt{houseA}}$ = **sum**(**max**(cosine) **group by** *Visual Preference*)

For query "I'm looking for a house with ***wooden floor, big windows*** *and* ***modern decoration****"*

Search

Scoring

House B

TEXT EMBEDDINGS

Wooden Floor `[0.1, 0.4, 0.6, 0.1]`

Big Windows `[0.2, 0.5, 0.7, 0.3]`

Modern Decoration `[0.1, 0.1, 0.6, 0.8]`

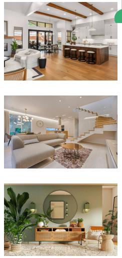
IMAGE EMBEDDINGS

`[0.2, 0.4, 0.7, 0.3]`

`[0.3, 0.5, 0.1, 0.1]`

`[0.6, 0.3, 0.2, 0.8]`

IMAGES



Pairwise
cosine similarity

$$\text{IMAGE\_SCORE}_{\text{houseB}} = \textbf{sum}(\textbf{max}(\text{cosine})\ \textbf{group by}\ \textit{Visual Preference})$$

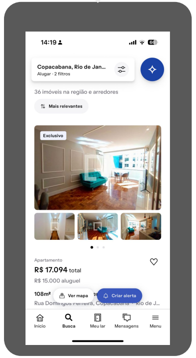
For query "I'm looking for a house with ***wooden floor, big windows*** *and* ***modern decoration****"*

Search
Scoring

$\text{SCORE}_{houseA} = \textbf{combine}(\text{IMAGE\_SCORE}_{houseA}, \text{HEAVY\_RANKER\_SCORE}_{houseA}, \ldots)$

$\text{SCORE}_{houseB} = \textbf{combine}(\text{IMAGE\_SCORE}_{houseB}, \text{HEAVY\_RANKER\_SCORE}_{houseB}, \ldots)$

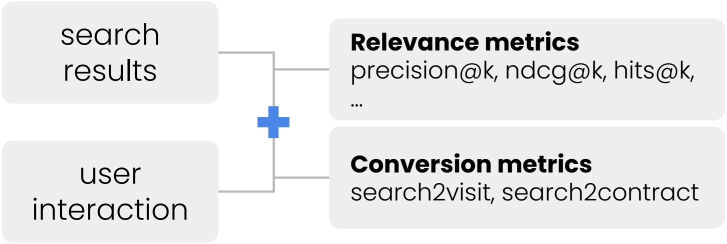
$\text{RESULTS} = \textbf{topK}(\textbf{rank}(\text{SCORE}_{houseA}, \text{SCORE}_{houseB}))$



But....is it good at all?

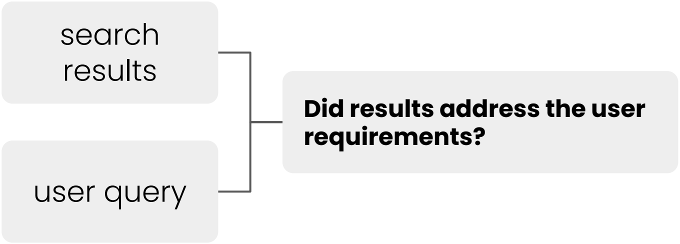
Search
Evaluation

Traditional



Overall health of the system

LLM-as-a-judge



Results (multimodal) explainability

Search

What we did	Future and WIP

• Similarity-based image sampling
• Different similarity aggregation strategies
  ○ AVG(SUM(cosine))
  ○ AVG(MAX(cosine))
  ○ MAX(MAX(cosine))
	• Hybrid Search: full-text + vector search
	• House-level vs. image-level embeddings
	• Heavy ranker with image features
	• Visual features as filter to reduce the search space

Thank You Team

We are hiring

Come join our team!

https://carreiras.quintoandar.com.br/?lang=en


Grupo
QuintoAndar