

AI Engineer Technical Challenge – Real Estate Website Content Generator

Problem Statement

You have been hired by a real estate company that manages hundreds of property listings across different cities and regions. The company wants to automate the creation of **high-quality, SEO-optimized, multilingual content** for property listing pages on their website.

Your task is to build an **AI-powered system** that generates all written content for these listing pages based on structured property data and a **fixed content structure**, with each section **tagged by specific HTML elements** so it can later be integrated dynamically into a website template.

Input

Structured property data in JSON format:

```
{
  "title": "T3 apartment in Lisbon",
  "location": {
    "city": "Lisbon",
    "neighborhood": "Campo de Ourique"
  },
  "features": {
    "bedrooms": 3,
    "bathrooms": 2,
    "area_sqm": 120,
    "balcony": true,
    "parking": false,
    "elevator": true,
    "floor": 2,
    "year_built": 2005
  },
  "price": 650000,
  "listing_type": "sale",
  "language": "en"
}
```

Website Content Structure (Text + HTML Tags)

Your system must generate all output in **plain text**, but each section should be **clearly marked with standard HTML tags**, as shown below. This allows easy templating without needing to render full HTML pages.

1. **<title>** – Page Title

(Appears in browser tab and search engine results – max 60 characters)

Example:

```
<title>T3 Apartment for Sale in Campo de Ourique, Lisbon</title>
```

2. **<meta name="description">** – Meta Description

(SEO snippet – max 155 characters)

Example:

```
<meta name="description" content="Spacious 3-bedroom apartment in Lisbon with balcony and elevator, located in Campo de Ourique. Ideal for families.">
```

3. **<h1>** – Headline

(Main visible headline on the page)

Example:

```
<h1>Modern T3 Apartment with Balcony in Campo de Ourique, Lisbon</h1>
```

4. **<section id="description">** – Full Property Description

(Rich, engaging paragraph with all key features. Use <p> inside. - between 500-700 characters)

Example:

```
<section id="description">
  <p>Located in the charming neighborhood of Campo de Ourique, this elegant T3 apartment offers 120 sqm of bright and spacious living.
```

Situated on the second floor of a well-maintained building with elevator access, the apartment features three bedrooms, two bathrooms, and a private balcony perfect for relaxing. Built in 2005, it combines modern amenities with timeless comfort. With a sale price of €650,000, this Lisbon property is ideal for families or professionals looking for a well-located home in the capital.</p></section>

5. <ul id="key-features"> – Key Features List

(3–5 short bullet points)

Example:

```
<ul id="key-features">
  <li>120 sqm of living space</li>
  <li>3 bedrooms and 2 bathrooms</li>
  <li>Private balcony</li>
  <li>Elevator access</li>
  <li>Located in Campo de Ourique, Lisbon</li>
</ul>
```

6. <section id="neighborhood"> – Neighborhood Summary

(One paragraph with lifestyle and area information)

Example:

```
<section id="neighborhood">
  <p>Campo de Ourique is one of Lisbon's most desirable
  neighborhoods, known for its vibrant cafés, green parks, and
  excellent schools. With a strong local community and easy access to
  the city center, it offers the perfect blend of charm and
  convenience.</p>
</section>
```

7. <p class="call-to-action"> – Call to Action

(Short closing line encouraging the user to act)

Example:

```
<p class="call-to-action">Don't miss this opportunity—schedule your viewing today and discover your new home in Lisbon.</p>
```

Output examples

- For clarity, we also send attached a html file with a **well-structured example** of the output that needs to be generated to be used only as a reference
 - The HTML tags **must be included exactly as shown**, for each corresponding section.
 - You are **not expected to render HTML** or create a frontend — just output 2-3 different examples with well-tagged text for dynamic insertion into an HTML template.
-

Requirements

Your solution must:

1. **Generate all 7 content sections**
 - Output plain text, but **each section must be wrapped with its required HTML tag** (as above).
 - Maintain the structure and order exactly.
2. **Support English and Portuguese (Portugal)**
 - Switch language using the **language** field in input.
 - Localize grammar, vocabulary, and idiomatic expressions accordingly.
3. **Ensure SEO Keyword Enrichment**
 - Each content section must **naturally include keywords** related to the location, property type, and listing intent.
 - Do not sacrifice readability or fluency.
4. **Handle Missing or Incomplete Data**
 - For example, if **elevator** or **year_built** is missing, generate content without awkward gaps or assumptions.
5. **Interface**

- Accept JSON as input and output structured text with the HTML-tagged content.
- Can be a CLI tool or REST API.

6. Documentation

- Include a README with:
 - Setup instructions
 - Example inputs/outputs
 - Any services or APIs used (e.g., OpenAI)
 - Assumptions and limitations

SEO Guidelines

DO:

- Include key **searchable phrases** like:
 - “apartment for sale in Lisbon”
 - “T3 apartment in Campo de Ourique”
 - “real estate in Portugal”
- Use **property-type and location combinations** in a natural tone.
- Include calls to action like “Schedule a visit” or “Don’t miss this opportunity”.

DON'T:

- Stuff keywords.
- Use repetitive or robotic language.
- Incorrect or misleading information

Objective of the Assessment

This is not a “perfect code” challenge — **we are not evaluating your ability to deliver polished production-ready software.**

Instead, the goal is to assess your ability to:

- Design **logical, structured and scalable solutions** to real-world AI problems
 - Think critically about trade-offs, limitations, and extensions
 - Communicate your approach clearly and defend your choices
-

Interview Format (1h30)

You do **not need to submit code** in advance. Instead, the challenge will be discussed in a **90-minute technical interview**, split into two parts:







1. **Solution Walkthrough (45 minutes)**

You will present your approach, architecture, assumptions, tools/models chosen, and how your system addresses each part of the challenge.







2. **Deep Dive & Discussion (45 minutes)**

We'll ask follow-up questions to explore your reasoning in more depth, review edge cases, and optionally explore extensions or trade-offs.

Bonus Points

-  Tone customization (formal, friendly, luxury, investor-focused)
 -  Add more languages (e.g., Spanish, French)
 -  Optionally support regional SEO tweaks (e.g., different keywords for UK vs PT)
 -  Grammar or readability evaluation tools
 -  Interactive testing UI (web-based or notebook)
 -  Evaluation framework to assess content quality and guideline compliance
-

Evaluation Criteria

-  Structure compliance (all HTML tags included, correct hierarchy)
-  Language fluency and SEO effectiveness
-  Multilingual adaptability
-  Code design and documentation
-  Robustness against input variation
-  Extra features (if included)

Estimated Time

Expect ~6-8 hours of focused work
