# Task Description: Real Estate Website Scraper Development for we-net.ch

## Overview

We require your expertise for a project involving the development of a scraper tool designed to extract and structure data from real estate websites. The initial phase will focus on scraping the following websites for properties which are on sale:

[Homegate.ch](http://homegate.ch/)

## Task Details

1. \*\*Scraper Tool Development:\*\* Utilize your expertise in web scraping tools and programming languages to develop a robust and efficient scraper for https://www.comparis.ch/. The scraper should be optimized for speed, accuracy, and reliability.

2. \*\*Data Extraction:\*\* The scraper should be able to collect comprehensive data from the website, including but not limited to the specific columns provided below.

3. \*\*Scraping Frequency:\*\* Initially, the scraper can be run manually for testing and optimization purposes. However, once finalized, it should be configured to run automatically every 5 minutes or at a predefined interval, ensuring a constant flow of up-to-date data.

4. \*\*Testing & Troubleshooting:\*\* Test the scraper for efficiency, and troubleshoot any issues that might occur during the data extraction process.

5. \*\*Scalability:\*\* Ensure the scraper is scalable and adaptable

## Deliverable: Structured Data Spreadsheet

We require the final data to be delivered in a structured and accessible spreadsheet format, with the following specific columns (always scrape as much datapoints as possible):

- Name Seller

- Phone number

- Email Address

- Property Type

- Rooms

- Built Year

- Property Address

- Listing Status

- Price

- Property features

- Property description

- Local amenities

- Website/property URL

The spreadsheet should be well-organized, with clearly labeled columns, and data should be formatted correctly for easy analysis and interpretation.

## Ethics & Legalities

Ensure that the data extraction respects the terms and conditions of the targeted website. The scraping practices should not infringe on privacy rights or breach any terms and conditions set by the website. All activities should remain within the boundaries of legal and ethical standards.