



# How to get clean web data for chatbots and LLMs

by František Nesveda



**Have you tried **chatting** with a website?**

- Model 3 Owner's Manual
- Overview
- Opening and Closing
- Storage Areas
- Seating and Safety Restraints
- Connectivity
- Driving
- Autopilot

About Autopilot

### Autopilot Features

Traffic Light and Stop Sign Control

Full Self-Driving (Supervised)

Autopark

Summon

Limitations and Warnings

- Active Safety Features

- Dashcam, Sentry, and Security

- Climate

- Navigation and Entertainment

- Charging and Energy Consumption

- Maintenance

- Specifications

# Autopilot Features

This topic describes how to enable and use the following driver assistance features.

#### NOTE

*Depending on market region, vehicle configuration, options purchased, and software version, your vehicle may not be equipped with the following features, or a feature may not operate exactly as described.*

- **Traffic-Aware Cruise Control:** Like traditional cruise control, Traffic-Aware Cruise Control maintains a set driving speed. However, Traffic-Aware Cruise Control also slows down or accelerates Model 3 as needed to maintain the following distance from the vehicle in front of you. While Traffic-Aware Cruise Control is engaged, you are still responsible for steering Model 3 (see [Traffic-Aware Cruise Control](#)).
- **Autosteer:** Like Traffic-Aware Cruise Control, Autosteer maintains a set speed (if there is not a vehicle in front of you) or a set following distance (if there is a vehicle in front of you). In addition, Autosteer detects lane markings, road edges, and the presence of vehicles and objects to intelligently keep Model 3 in its driving lane (see [Autosteer](#)).

#### NOTE

Autosteer is a BETA feature.

- **Auto Lane Change:** When you engage a turn signal while Autosteer is active, Auto Lane Change moves Model 3 into the adjacent lane in the direction indicated by the turn signal (see [Auto Lane Change](#)).
- **Navigate on Autopilot:** Navigate on Autopilot builds on the features of Traffic-Aware Cruise Control and Autosteer. While Autosteer is active, Navigate on Autopilot allows Model 3 to suggest and, if configured,

```
import cheerio from 'cheerio';
import OpenAI from 'openai';

const QUESTION = 'How to turn on Autopilot?';
const URL = 'https://www.tesla.com/ownersmanual/model3/en_us/GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B.html';

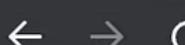
const response = await fetch(URL);

const $ = cheerio.load(response.body);
const body = $('body');

const openai = new OpenAI();
const chatCompletion = await openai.chat.completions.create({
  model: 'gpt-4o',
  messages: [
    {
      role: 'user',
      content: `Answer the question: "${QUESTION}" using information from the following HTML: ${body}`,
    },
    temperature: 0,
  ],
});

console.log(chatCompletion.choices);
```



[Home](#)[BMW ConnectedDrive](#)[Electromobility](#)[Models](#)[BMW USA](#)[Contact Us](#) Search Frequently Asked Questions[Home > Article](#)

# Article.

118715

## How can I restart or "reboot" the BMW iDrive system in my vehicle?

To restart the BMW iDrive system in your BMW, press and hold the volume control button on the audio control unit for at least 30 seconds. Please note: Only restart the BMW iDrive system in your vehicle when the vehicle is safely parked. Your vehicle must be in the "radio-ready" state (BMW iDrive active). Some functions of the BMW iDrive operating system will not be available or only partially available during the restart. Please allow the system to fully restart before driving the vehicle.

[ENTERTAINMENT & CONNECTIVITY](#)[TECHNOLOGIES & INNOVATIONS](#)[MODELS & EQUIPMENT](#)

Was this article helpful?

19    26

## Recommended articles.

19967

## How do I activate split screen mode in the control display of my BMW?

To activate split screen mode in the control display of your BMW with Operating System 7, select: "NAV" > "Split screen". In your BMW with iDrive operating system with main menu...

[> Show full article](#)

5476

Can I display the calendar of the day in my BMW i3?

```
const URL = 'https://faq.bmwusa.com/s/article/BMW-iDrive-Displays-and-operation-Reset-Restart-2WPKA';  
const QUESTION = 'How to restart iDrive?';
```

```
message: {  
  role: 'assistant',  
  content: 'The provided HTML does not contain any information about how to restart iDrive.'  
}
```



**ONE DOES NOT SIMPLY**

**CHAT WITH A WEBSITE**

# How to get clean web data for chatbots and LLMs

**František Nesveda**

Full-stack engineer & Team leader



# 1. Reliably **fetching** page content.

- 1. Reliably fetching page content.**
- 2. Parsing HTML for LLMs.**

1. Reliably **fetching** page content.
2. **Parsing** HTML for LLMs.
3. **Crawling** all the pages you need.

1. Reliably **fetching** page content.
2. **Parsing** HTML for LLMs.
3. **Crawling** all the pages you need.
4. Creating a **chatbot** with LangChain.



# Website Content Crawler

[apify/website-content-crawler](#)



Developed by



Apify

AI   Developer tools

**35.7k** users   **8.7M** runs

AI BOT

# Automate conversations at the speed of AI



INTERCOM

The screenshot shows a conversational interface between a user and an AI bot named Fin. The user asks, "Is there a fee if I change the dates?" The bot responds with, "No, there's no charge to change the date of your reservation, just remember that it can only be changed up to seven days before it is due to begin." Below this, a "Source" link points to "Modifying your reservation >". The user then asks, "Did that help?", and the bot replies with an "Answer" button and an information icon.

Fin

Ask our bot a question

The team can help if needed

Hi! How can I help?

Fin

AI Bot

Is there a fee if I change the dates?

No, there's no charge to change the date of your reservation, just remember that it can only be changed up to seven days before it is due to begin.

Source

Modifying your reservation >

Answer

Did that help?

# Reliably fetching page content.

```
curl https://www.tesla.com/ownersmanual/model3/...
```

curl: (56) HTTP/2 stream 1 was reset

**We must be **human-like**.**

```
{  
  'user-agent': 'curl/7.64.1',  
  accept: '*/*'  
}
```

```
{  
    'sec-ch-ua': '"Chromium";v="92", " Not A;Brand";v="99", "Google Chrome";v="92"' ,  
    'sec-ch-ua-mobile': '?0' ,  
    'upgrade-insecure-requests': '1' ,  
    'user-agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like  
Gecko) Chrome/92.0.4515.107 Safari/537.36' ,  
    'accept': 'text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image  
/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.9' ,  
    'sec-fetch-site': 'same-site' ,  
    'sec-fetch-mode': 'navigate' ,  
    'sec-fetch-user': '?1' ,  
    'sec-fetch-dest': 'document' ,  
    'accept-encoding': 'gzip, deflate, br' ,  
    'accept-language': 'en-GB,en-US;q=0.9,en;q=0.8'  
}
```

**HTTP 3**

**Headers**

**TLS versions**



**ECDH Curves**

**Signature  
Algorithms**

**Ciphers**

[README](#)[MIT license](#)

# curl-impersonate

[Build and test](#)

passing

[Docker images](#)

passing

A special build of [curl](#) that can impersonate the four major browsers: Chrome, Edge, Safari & Firefox. curl-impersonate is able to perform TLS and HTTP handshakes that are identical to that of a real browser.

curl-impersonate can be used either as a command line tool, similar to the regular curl, or as a library that can be integrated instead of the regular libcurl. See [Usage](#) below.

# impit | browser impersonation made simple

impit is a `rust` library that allows you to impersonate a browser and make requests to websites. It is built on top of `reqwest`, `rustls` and `tokio` and supports HTTP/1.1, HTTP/2, and HTTP/3.

The library provides a simple API for making requests to websites, and it also allows you to customize the request headers, use proxies, custom timeouts and more.

```
use impit::impit::Impit;
use impit::emulation::Browser;

#[tokio::main]
async fn main() {
    let mut impit = Impit::builder()
        .with_browser(Browser::Firefox)
        .with_http3()
        .build();

    let response = impit.get(String::from("https://example.com"), None).await;
```

```
import cheerio from 'cheerio';
import { Impit } from 'impit';
import OpenAI from 'openai';

const QUESTION = 'How to turn on Autopilot?';
const URL = 'https://www.tesla.com/ownersmanual/model3/en_us/GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B.html';

const impit = new Impit({ browser: 'Chrome' });
const response = await impit.fetch(URL); ←

const $ = cheerio.load(response.body);
const body = $('body');

const openai = new OpenAI();
const chatCompletion = await openai.chat.completions.create({
  model: 'gpt-4o',
  messages: [
    {
      role: 'user',
      content: `Answer the question: "${QUESTION}" using information from the following HTML: ${body}`,
    },
    temperature: 0,
  ],
});

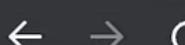
console.log(chatCompletion.choices);
```



```
message: {  
  role: 'assistant',  
  content: 'To turn on Autopilot, follow these steps: 1. Make sure your Tesla vehicle...'}  
}
```

**Are we human-like yet?**

**No. Browsers also execute JavaScript.**

[Home](#)[BMW ConnectedDrive](#)[Electromobility](#)[Models](#)[BMW USA](#)[Contact Us](#) Search Frequently Asked Questions[Home > Article](#)

# Article.

118715

## How can I restart or "reboot" the BMW iDrive system in my vehicle?

To restart the BMW iDrive system in your BMW, press and hold the volume control button on the audio control unit for at least 30 seconds. Please note: Only restart the BMW iDrive system in your vehicle when the vehicle is safely parked. Your vehicle must be in the "radio-ready" state (BMW iDrive active). Some functions of the BMW iDrive operating system will not be available or only partially available during the restart. Please allow the system to fully restart before driving the vehicle.

[ENTERTAINMENT & CONNECTIVITY](#)[TECHNOLOGIES & INNOVATIONS](#)[MODELS & EQUIPMENT](#)

Was this article helpful?

19    26

## Recommended articles.

19967

## How do I activate split screen mode in the control display of my BMW?

To activate split screen mode in the control display of your BMW with Operating System 7, select: "NAV" > "Split screen". In your BMW with iDrive operating system with main menu...

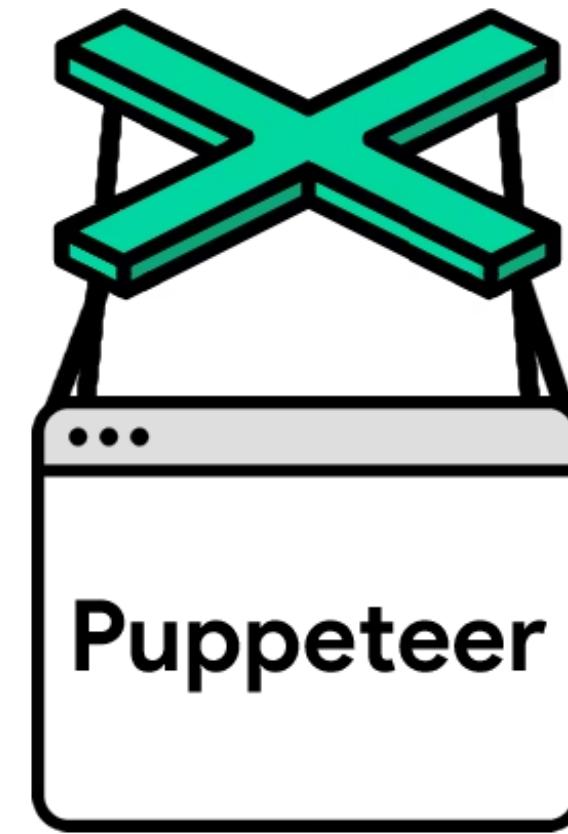
[> Show full article](#)

5476

Can I display the calendar of the day in my BMW i3?



# Selenium



# Playwright

```
import playwright from 'playwright';

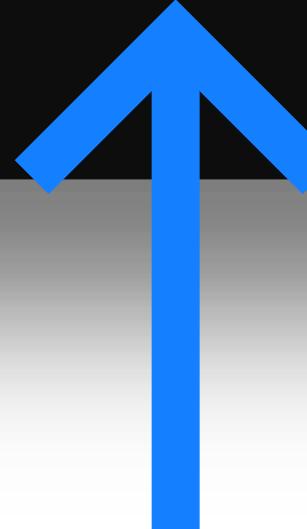
const browser = await playwright.chromium.launch();
const page = await browser.newPage();

await page.goto(URL);

const html = await page.content();
```

# Don't forget to **wait** for content to load.

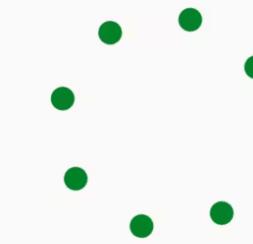
```
await page.goto(URL, { waitUntil: 'networkidle' }));
```





```
message: {  
  role: 'assistant',  
  content: 'To restart the BMW iDrive system in your vehicle, press and hold...'  
}
```

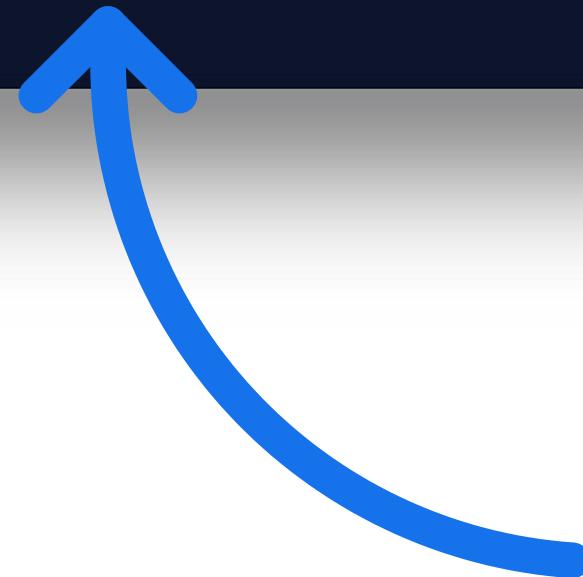
**Prepare to fight anti-bot software.**



Verifying...



Mozilla/5.0 (Macintosh; Intel Mac OS X 10\_15\_7) AppleWebKit/537.36 (KHTML, like Gecko)  
HeadlessChrome/94.0.4595.0 Safari/537.36



```
{  
  fingerprint: {  
    userAgent: 'Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko)  
Chrome/92.0.4515.131 Safari/537.36',  
    cookiesEnabled: true,  
    timezone: 'America/Toronto',  
    timezoneOffset: 300,  
    audioCodecs: { ogg: 'probably', mp3: 'probably', wav: 'probably', m4a: 'maybe', aac: 'probably' },  
    videoCodecs: { ogg: 'probably', h264: 'probably', webm: 'probably' },  
    battery: true,  
    videoCard: [ 'Intel Inc.', 'Intel(R) Iris(TM) Plus Graphics 655' ],  
    deviceMemory: 8,  
    productSub: '20030107',  
    hardwareConcurrency: 16,  
    vendor: 'Google Inc.',  
    multimediaDevices: { speakers: 1, micros: 1, webcams: 1 },  
    platform: 'MacIntel',  
    pluginsSupport: true,  
    plugins: [ /* ... */ ],  
    mimeTypes: [ /* ... */ ],  
    screenResolution: [ 2560, 1440 ],  
    availableScreenResolution: [ 2560, 1361 ],  
    colorDepth: 24,  
    touchSupport: { maxTouchPoints: 0, touchEvent: false, touchStart: false },  
    languages: [ 'en-US' ]  
},  
}
```

**puppeteer-extra-plugin-stealth**



**Playwright**

**selenium-stealth**



**crawlee**

# What about **web scraping APIs**?

Easy to start with, but costly at scale.

# Parsing HTML for LLMs.

- ✓ Model 3 Owner's Manual
- ✓ Overview
- ✓ Opening and Closing
- ✓ Storage Areas
- ✓ Seating and Safety Restraints
- ✓ Connectivity
- ✓ Driving
- ✓ Autopilot

### About Autopilot

#### Autopilot Features

Traffic Light and Stop Sign Control

Full Self-Driving (Beta)

Autopark

Summon

Smart Summon

#### Limitations and Warnings

✓ Active Safety Features

✓ Dashcam, Sentry, and Security

✓ Climate

✓ Navigation and Entertainment

✓ Charging and Energy Consumption

.. Maintenance

# About Autopilot

Autopilot is a suite of advanced driver assistance features that are intended to make driving safer and less stressful. None of these features make Model 3 fully autonomous or replace you as the driver. Autopilot features come standard with all new Tesla vehicles.



#### Note

Depending on market region, vehicle configuration, options purchased, and software version, your vehicle may not be equipped with all features listed below, or a feature may not operate as described.

Basic Autopilot includes Traffic-Aware Cruise Control and Autosteer.

- **Traffic-Aware Cruise Control:** Maintains your speed and an adjustable following distance from the vehicle in front of you, if there is one (see [Traffic-Aware Cruise Control](#)).
- **Autosteer:** Maintains your speed and distance from a leading vehicle while also intelligently keeping Model 3 in its lane (see [Autosteer](#)).



#### Warning

Basic Autopilot is a hands-on feature. Keep your hands on the steering wheel at all times and be mindful of road conditions, surrounding traffic, and other road users (such as pedestrians and cyclists). Always be prepared to take immediate action. Failure to follow these instructions could cause damage, serious injury, or death.

Enhanced Autopilot includes additional features. Enhanced Autopilot Features are designed to further

```
<html lang="en-us">
  ▶ <head> ...
  ▼ <body class="site-body site-js-body tds-fullscreen-ios">
    ▼ <div class="tds-shell" id="GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B"> grid
      ▶ <div class="tds-shell-masthead tds-shell-masthead--is_sticky"> ... </div> flex
        <div id="search-results" class="search-results" style="display: none;"></div>
      ▼ <div class="tds-shell-content">
        ▼ <div class="tds-layout tds-layout-2col-static_sidebar--large tds-layout-main--right">
          ▼ <div class="tds-layout-item tds-layout-main">
            ▼ <main role="main" class="tds-shell-content" id="main-content">
              ▼ <article role="article" aria-labelledby="GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B__GUID-F473BEE0-E655-44AA-AF72-8E4CB0A0F033">
                <h1 class="title topicTitle1" id="GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B__GUID-F473BEE0-E655-44AA-AF72-8E4CB0A0F033">About
                  Autopilot</h1>
                ▼ <div class="body conbody">
                  ▶ <p class="p"> ...
                  ▶ <div class="tds-status_msg tds-card tds-card--dense" role="status"> ... </div> flex
                    <p class="p">Basic Autopilot includes Traffic-Aware Cruise Control and Autosteer.</p>
                  ▶ <ul class="ul tds-list tds-list--unordered" id="GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B__GUID-C95D8F8B-437B-487F-9959-E445E39A0C1
                      8"> ... </ul>
                  ▶ <div class="tds-status_msg tds-card tds-card--dense" role="status"> ... </div> flex
                  ▶ <p class="p"> ...
                  ▶ <ul class="ul tds-list tds-list--unordered" id="GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B__GUID-10FD291D-93B1-43A3-ADD6-C0AA489767D
                      A"> ... </ul>
                  ▶ <div class="tds-status_msg tds-card tds-card--dense" role="status"> ... </div> flex
                    <p class="p">Full Self-Driving Capability includes:</p>
                  ▶ <ul class="ul tds-list tds-list--unordered" id="GUID-101D1BF5-52D2-469A-A57D-E7230BBEE94B__GUID-2ABEB5BD-8601-4F6D-9E7C-6B334816BC2
                      B"> ... </ul>
                  ▶ <div class="tds-status_msg tds-card tds-card--dense" role="status"> ... </div> flex
                  ▶ <p class="p"> ...
                  ▶ <p class="p"> ...
                  ▶ <p class="p"> ...
                  ▶ <div class="tds-status_msg tds-card tds-card--dense" role="status"> ... </div> flex
                  ▶ <p class="p"> ...
```

**Start with [@mozilla/readability](#)**

Enhanced Autopilot includes additional features. Enhanced Autopilot Features are designed to further reduce driver workload and make common actions, such as changing lanes or parking, easier.

- Auto Lane Change: Moves Model 3 into an adjacent lane when you engage the turn signal and Autosteer is active (see [Auto Lane Change](#)).
- Navigate on Autopilot: Actively guides Model 3 from a highway's on-ramp to off-ramp, including performing lane changes, navigating interchanges, automatically engaging the turn signal, and taking the correct exit (see [Navigate on Autopilot](#)).
- Autopark: Parks Model 3, either parallel or perpendicularly (see [Autopark](#)).
- Summon: Moves Model 3 forward or backward, even while you're outside the vehicle. This is useful for parking in tight parking spots (see [Summon](#)).
- Smart Summon: Moves Model 3 out of a parking space and through more complex environments, maneuvering around obstacles and other vehicles, to meet you or go to a predetermined target (see [Smart Summon](#)).

**Sadly, Readability isn't 100% reliable.**

**Keep noise out of your embeddings.**

# noise

- ✓ Model 3 Owner's Manual
- ✓ Overview
- ✓ Opening and Closing
- ✓ Storage Areas
- ✓ Seating and Safety Restraints
- ✓ Connectivity
- ✓ Driving
- ✓ Autopilot

## About Autopilot

### Autopilot Features

Traffic Light and Stop Sign Control

Full Self-Driving (Beta)

Autopark

# About Autopilot

Autopilot is a suite of advanced driver assistance features designed to make driving less stressful. None of these features make Model 3 fully autonomous. These features come standard with all new Tesla vehicles.



### Note

Depending on market region, vehicle configuration, and options, your vehicle may not be equipped with all features.

Basic Autopilot includes Traffic-Aware Cruise Control:

- **Traffic-Aware Cruise Control:** Maintains your speed and distance from the vehicle in front of you, if there is one (see [Traffic-Aware Cruise Control](#))
- **Autosteer:** Maintains your speed and distance from the vehicle in front of you, if there is one (see [Autosteer](#))

- Seating
- Storage
- Driving
- ...

**Autopilot is a suite  
of advanced ...**

- Seating
- Storage
- Driving
- ...

**When you open a  
door ...**

- Seating
- Storage
- Driving
- ...

**Model 3 is  
designed to ...**

# **Remove shared lines (Count-min Sketch)**

**Autopilot is a suite  
of advanced ...**

**When you open a  
door ...**

**Model 3 is  
designed to ...**

# Close **popups** to prevent wrong extractions:

**idcac-playwright** TS

0.1.3 • [Public](#) • Published 6 months ago

[Readme](#) [Code](#) Beta [0 Dependencies](#)

## I Don't Care About Cookies for Playwright/Puppeteer

This package contains the **I don't care about cookies** browser extension compiled for use with Playwright or Puppeteer.

# Convert to markdown

```
1 # About Autopilot
2
3 Autopilot refers to a suite of advanced driver assistance
4 features that are intended to make driving safer and less
5 stressful. None of these features make Model Y fully
6 autonomous or replace you as the driver. Autopilot features
7 come standard with all new Tesla vehicles.
8
9 Note
10
11 Depending on market region, vehicle configuration, options
12 purchased, and software version, your vehicle may not be
13 equipped with the following features, or a feature may not
14 operate exactly as described.
15
16 Autopilot includes Traffic-Aware Cruise Control and
17 Autosteer.
18
19 * **Traffic-Aware Cruise Control**: Maintains your speed
20 and an adjustable following distance from the vehicle in
21 front of you, if there is one (see [Traffic-Aware Cruise
22 Control](https://www.tesla.com/ownersmanual/model3/en\_us/GUID-20F2262F-CDF6-408E-A752-2AD9B0CC2FD6.html#GUID-842DBF70-E101-4250-8CA5-DCE6135D7149)).
```

12 \* \*\*Autosteer\*\*: Maintains your speed and distance from a leading vehicle while also intelligently keeping Model 3 in its lane (see [Autosteer]([https://www.tesla.com/ownersmanual/model3/en\\_us/GUID-20F2262F-CDF6-408E-A752-2AD9B0CC2FD6.html#GUID-E006C4C8-C0CB-4819-BD02-AE34C936025F](https://www.tesla.com/ownersmanual/model3/en_us/GUID-20F2262F-CDF6-408E-A752-2AD9B0CC2FD6.html#GUID-E006C4C8-C0CB-4819-BD02-AE34C936025F)))).

**Crawling all the pages you need.**

**Explore like a human. Ignore sitemaps.**

**Link parsing**

**State  
management**

**Queueing &  
Uniqueness**

**Crawling**

**Scope &  
Redirects**

**URL  
resolution**

**Concurrency**



Scrapy

Use a crawling library.



crawlee

**To crawl at scale, you will need proxies.**

# Datacenter proxies:

Fast, cheap, high bandwidth, often blocked.

# **Residential proxies:**

Flaky speed, expensive, rarely blocked.

# Golden rule of proxy selection:

Always use **datacenter** proxies. Unless they are blocked. Only then use residential.

# Creating a **chatbot** with LangChain.

```
import { CheerioCrawler, Dataset } from 'crawlee';

const URL = 'https://www.tesla.com/ownersmanual/model3/en_us/';

const crawler = new CheerioCrawler({
  requestHandler: async ({ body, request, enqueueLinks }) => {
    console.log('Parsing', request.url)

    await enqueueLinks({
      globs: [`#${URL}/**`]
    })
  }
});

await crawler.run([URL]);
```

```
INFO CheerioCrawler: Starting the crawler.  
Parsing https://www.tesla.com/ownersmanual/model3/en_us/  
Parsing https://www.tesla.com/ownersmanual/model3/en_us/GUID-6C6C3944-9674-4E81-  
Parsing https://www.tesla.com/ownersmanual/model3/en_us/GUID-EA1715B0-A3A6-454E-  
Parsing https://www.tesla.com/ownersmanual/model3/en_us/GUID-80B80D48-E3A9-4857-  
Parsing https://www.tesla.com/ownersmanual/model3/en_us/GUID-59E1B254-DB74-4178-  
Parsing https://www.tesla.com/ownersmanual/model3/en_us/GUID-A5F2B9D0-E7C8-40F5-  
...
```

```
import { Readability } from '@mozilla/readability';
import { JSDOM } from 'jsdom';
import TurndownService from 'turndown';

// ...

const dom = new JSDOM(body, { url: request.loadedUrl });
const reader = new Readability(dom.window.document, {
    charThreshold: 200,
});

const { content } = reader.parse();
const markdown = turndownService.turndown(content);
```

```
import { CSVLoader } from '@langchain/community/document_loaders/fs/csv';

await Dataset.exportToCSV('tesla-docs');

const loader = new CSVLoader(
  './tesla-docs.csv',
  'markdown',
);

const docs = await loader.load();
```

```
import { OpenAIEmbeddings } from '@langchain/openai';
import { RecursiveCharacterTextSplitter } from '@langchain/textsplitters';
import { MemoryVectorStore } from 'langchain/vectorstores/memory';

const textSplitter = new RecursiveCharacterTextSplitter({
    chunkSize: 500,
    chunkOverlap: 50,
});

const splitDocs = await textSplitter.splitDocuments(docs);

const embeddings = new OpenAIEmbeddings();
const vectorStore = await MemoryVectorStore.fromDocuments(
    splitDocs,
    embeddings,
);
```

```
const model = new ChatOpenAI({
  model: 'gpt-4o',
  apiKey: OPENAI_API_KEY,
  temperature: 0,
});

const questionAnsweringPrompt = ChatPromptTemplate.fromMessages([
  ['system', 'Answer the user\'s questions based on the below context:\n\n{context}'],
  ['human', '{input}'],
]);

const combineDocsChain = await createStuffDocumentsChain({
  llm: model,
  prompt: questionAnsweringPrompt,
});

const chain = await createRetrievalChain({
  retriever: vectorStore.asRetriever(),
  combineDocsChain,
});

const QUESTION = 'What is Autopilot and how to use it?';
const res = await chain.invoke({ input: QUESTION });
console.log(res.answer);
```

```
{  
    text: 'Autopilot is a driver assistance feature in Tesla vehicles that helps with steering,  
accelerating, and braking. It is designed to assist the driver and make driving easier and more  
convenient.\n' +  
    '\n' +  
    'To turn on Autopilot, follow these steps:\n' +  
    '\n' +  
    '1. Make sure your Tesla vehicle is in Park and you are safely parked.\n' +  
    "2. On the touchscreen display, tap on the 'Controls' icon located at the bottom of the screen.\n"  
    +  
    "3. In the Controls menu, select 'Autopilot'.\n" +  
    "4. In the Autopilot menu, you will see various options. One of the options is 'Enable at Start of  
Every Trip'. If this option is turned off, you will need to manually enable Autopilot for each trip. If  
it is turned on, Autopilot will be automatically enabled at the start of every trip.\n" +  
    "5. If you want to enable Autopilot for the current trip, touch the 'Navigate on Autopilot' button  
located above the turn-by-turn directions. This will engage Autopilot whenever you engage Autosteer.\n"  
    +  
    "6. Once Autopilot is enabled, you will see the Autopilot icon on the touchscreen turn blue,  
indicating that it is active.\n' +  
    '\n' +  
    'Please note that the availability and functionality of Autopilot features may vary depending on  
your market region and the specific Tesla model you have. It is important to always stay attentive and  
be ready to take control of the vehicle when using Autopilot.'  
}
```

```
import { PlaywrightCrawler } from 'crawlee';

const URL = 'https://faq.bmwusa.com/s/';

const crawler = new PlaywrightCrawler();
const body = page.content();

const QUESTION = 'Explain what iDrive is ' +
    'and provide step by step instructions to restarting it.';
```

```
INFO  PlaywrightCrawler: Starting the crawler.  
Parsing https://faq.bmwusa.com/s/  
Parsing https://faq.bmwusa.com/s/article/My-BMW-App-Remote-Services-Beschre...  
Parsing https://faq.bmwusa.com/s/article/BMW-tracking-new-vehicle-order-iTA...  
Parsing https://faq.bmwusa.com/s/article/BMW-iDrive-Displays-and-operation-...  
Parsing https://faq.bmwusa.com/s/article/Connected-Music-Spotify-Login-D9U1...  
Parsing https://faq.bmwusa.com/s/article/My-BMW-portal-Login-Login-Login-d...  
Parsing https://faq.bmwusa.com/s/article/Digital-Key-Key-Card-RFID-Activati...  
Parsing https://faq.bmwusa.com/s/article/My-BMW-App-Add-vehicle-Instruction...  
...
```

```
{  
    text: 'iDrive is a system used in BMW vehicles that controls various functions such as navigation, entertainment, and communication. It is operated through a central control knob and a display screen.\n' +  
        '\n' +  
        'To restart the iDrive system in your BMW, follow these step-by-step instructions:\n' +  
        '\n' +  
        '1. Ensure that your vehicle is safely parked and in the "radio-ready" state (i.e., the iDrive system is active).\n' +  
        '2. Locate the volume control button on the audio control unit. It is usually located near the central control knob.\n' +  
        '3. Press and hold the volume control button for at least 30 seconds. Make sure to hold it down continuously without releasing it.\n' +  
        '4. During this time, some functions of the iDrive system may not be available or may only be partially available. This is normal during the restart process.\n' +  
        '5. After holding the volume control button for 30 seconds, release it.\n' +  
        '6. Allow the iDrive system to fully restart. This may take a few moments.\n' +  
        '7. Once the system has fully restarted, you can use the iDrive functions as usual.\n' +  
        '\n' +  
        "Please note that these instructions are specific to restarting the iDrive system in a BMW vehicle.  
        If you have any concerns or difficulties, it is recommended to consult your vehicle's owner's manual or contact a BMW dealership for further assistance."  
}
```

 **APIFY**

# Website Content Crawler

[apify/website-content-crawler](#)

Developed by  **Apify**

 [AI](#) [Developer tools](#)

**35.7k** users **8.7M** runs





**Thank you.**



[github.com/apify/devworld2025](https://github.com/apify/devworld2025)

