

# URL

## Contents

- URL
- Selenium URL Loader

This covers how to load HTML documents from a list of URLs into a document format that we can use downstream.

```
from langchain.document_loaders import UnstructuredURLLoader
```

```
urls = [  
    "https://www.understandingwar.org/backgrounder/russian-offensive-campaign-  
assessment-february-8-2023",  
    "https://www.understandingwar.org/backgrounder/russian-offensive-campaign-  
assessment-february-9-2023"  
]
```

```
loader = UnstructuredURLLoader(urls=urls)
```

```
data = loader.load()
```

This covers how to load HTML documents from a list of URLs using the `SeleniumURLLoader`.

Using selenium allows us to load pages that require JavaScript to render.

## Setup

To use the `SeleniumURLLoader`, you will need to install `selenium` and `unstructured`.

[Skip to main content](#)

```
urls = [  
    "https://www.youtube.com/watch?v=dQw4w9WgXcQ",  
    "https://goo.gl/maps/NDSHwePEyaHMFgwh8"  
]
```

```
loader = SeleniumURLLoader(urls=urls)
```

```
data = loader.load()
```