

```
import requests
from bs4 import BeautifulSoup
from gensim.summarization import summarize
from gensim.summarization import keywords
from gensim.utils import simple_preprocess
```

In this section, I've imported the required libraries, requests for sending HTTP requests, BeautifulSoup for parsing HTML, and gensim for text summary and keyword extraction.

```
url = "https://www.zhihu.com/"
response = requests.get(url)
soup = BeautifulSoup(response.content, "html.parser")
```

Here I sent a GET request to the home page of Zhihu and got a response parsed by BeautifulSoup into a traversable object soup.

```
text = ""
for p in soup.find_all("p"):
    text += p.get_text()
```

I find all the p elements and merge their text into one long string text.

```
tokens = simple_preprocess(text)
```

Next, I use the simple_preprocess function to convert the text into a list, where each element is a processed word.

```
kw = keywords(text, ratio=0.1, lemmatize=True)
query = "+".join(kw)
```

Then, I used the keywords function in gensim to extract the keywords from the text and set ratio=0.1 to mean that the number of keywords returned did not exceed 10% of the total number of words in the text. Then, I concatenated these keywords into a string and used it as a query string for searching Zhihu.

```
search_url = f"https://www.zhihu.com/search?type=content&q={query}"
search_response = requests.get(search_url)
search_soup = BeautifulSoup(search_response.content, "html.parser")
```

I send a GET request to the URL and parse the returned HTML into a soup object.

```
results = search_soup.find_all("div", {"class": "SearchItem"})
for result in results:
    result_text = result.get_text()
    result_tokens = simple_preprocess(result_text)
    summary = summarize(result_text, ratio=0.2)
    print(f"URL: {result.a['href']}")
    print(f"Summary: {summary}\n")
```

I find all the search results and extract the text of each result using BeautifulSoup's `find_all` method. I processed the text into a group of words and used the `summarize` function in `gensim` to summarize it into a paragraph, setting `ratio=0.2` to indicate that the length of the output summary text did not exceed 20% of the length of the original text. Finally, I print the URL and summary text of the search results.