```
In [1]:  import requests
         from bs4 import BeautifulSoup
         from urllib.parse import urljoin

In [2]:  def web_crawler(seed_url, max_pages=10):
             visited = set()
             to_visit = [seed_url]

             while to_visit and len(visited) < max_pages:
                 url = to_visit.pop(0)     # BFS → queue
                 if url in visited:
                     continue

                 try:
                     response = requests.get(url, timeout=5)
                     if response.status_code != 200:
                         continue

                     visited.add(url)
                     print(f"\nCrawled: {url}")

                     soup = BeautifulSoup(response.text, "html.parser")

                     # Extract all links
                     for link in soup.find_all("a", href=True):
                         full_url = urljoin(url, link["href"])
                         if full_url not in visited and full_url.startswith("http"):
                             to_visit.append(full_url)

                 except Exception as e:
                     print(f"Error fetching {url}: {e}")

             return visited

In [9]:  seed = "https://www.youtube.com/"
         crawled_urls = web_crawler(seed, max_pages=15)
```

```
Crawled: https://www.youtube.com/

Crawled: https://www.youtube.com/about/

Crawled: https://www.youtube.com/about/press/

Crawled: https://www.youtube.com/about/copyright/

Crawled: https://www.youtube.com/t/contact_us/

Crawled: https://www.youtube.com/creators/

Crawled: https://www.youtube.com/ads/

Crawled: https://developers.google.com/youtube

Crawled: https://www.youtube.com/t/terms

Crawled: https://www.youtube.com/t/privacy

Crawled: https://www.youtube.com/about/policies/

Crawled: https://www.youtube.com/howyoutubeworks?utm_campaign=ytgen&utm_source=yt
hp&utm_medium=LeftNav&utm_content=txt&u=https%3A%2F%2Fwww.youtube.com%2Fhowyoutub
eworks%3Futm_source%3Dythp%26utm_medium%3DLeftNav%26utm_campaign%3Dytgen

Crawled: https://www.youtube.com/new

Crawled: https://www.youtube.com/about/#content

Crawled: https://www.youtube.com/howyoutubeworks/
```

In [ ]: