5회차 과제 - Web

Date	@2024년 7월 25일 오후 11:59

- HanKyungScraper.py
  - `__init__(self)`
    - webdriver.Chrome을 사용하여 브라우저 설정
  - `get_article_urls(self, start_date: datetime, end_date: datetime) -> List[str]`
    - 주어진 날짜 범위에 해당하는 경제 카테고리 기사 URL 수집
      1. 경제 카테고리 진입
      2. 경제 카테고리에서 보여주고 있는 기사의 URL만 수집 (전체 카테고리 X)

         - <div class=”box-module-inner economyDiv slick-slide slick-current slick-active” ... > 태그의 하위 태그들로부터 경제 카테고리에 해당하는 기사들의 URL을 수집할 수 있음
      3. 을 하나씩 순회하며 범위에 맞는 날짜의 경제 카테고리 기사 URL을 수집
         - 의 인덱스를 통해 해당 태그를 순회
         - 만약 다음 인덱스가 없는 경우 더보기 버튼을 클릭해서 다음 인덱스에 해당하는 태그를 불러옴
  - `url_to_data(self, url_list: List[str], output_path: str) -> None`
    - 수집한 기사 URL 목록으로부터 각 기사의 제목, 입력일자, 수정일자, 본문 등을 크롤링하여 JSON 형식으로 저장

- `close(self)`
  - 브라우저 종료