

SmartCrawler Pro 사용설명서

목차

1. [프로그램 소개](#)
 2. [기본 사용법](#)
 3. [고급 기능](#)
 4. [사이트별 크롤링 팁](#)
 5. [문제 해결](#)
-

프로그램 소개

SmartCrawler Pro는 웹사이트의 데이터를 쉽고 빠르게 수집할 수 있는 크롤링 도구입니다.

주요 기능

- 정적 웹페이지 크롤링 (requests 방식)
 - 동적 웹페이지 크롤링 (Selenium 방식)
 - 여러 페이지 자동 순회
 - 크롤링 결과 엑셀 저장
 - 사이트별 맞춤 크롤링
 - 실시간 진행률 표시
-

기본 사용법

1. 프로그램 시작

- 바탕화면의 "SmartCrawler Pro" 아이콘을 더블클릭
- 또는 시작 메뉴에서 실행

2. 크롤링 설정

1. **URL 입력:** 크롤링하고 싶은 웹사이트 주소 입력
2. **브라우저 모드 선택:**
3. **requests:** 빠른 정적 페이지용
4. **selenium:** 동적 페이지용 (JavaScript 필요시)
5. **페이지 수:** 몇 페이지까지 크롤링할지 설정
6. **크롤링 간격:** 페이지 간 대기 시간 (초)

3. 크롤링 실행

1. "크롤링 시작" 버튼 클릭
2. 진행률 바에서 진행 상황 확인
3. 완료되면 "결과 테이블" 탭에서 데이터 확인

4. 결과 저장

- "엑셀로 저장" 버튼 클릭
- 파일 이름과 저장 위치 선택
- Excel 파일 (.xlsx)로 저장

고급 기능

브라우저 모드 선택 가이드

requests 모드 (권장) - 빠른 속도 - 적은 자원 사용 - 정적 HTML 페이지에 적합 - 예: 뉴스 사이트, 블로그, 기본 쇼핑몰

selenium 모드 - 느린 속도 - 많은 자원 사용 - JavaScript가 필요한 동적 페이지에 적합 - 예: SPA (Single Page Application), React 사이트

크롤링 간격 설정

- 1-2초: 일반적인 사이트
- 3-5초: 큰 사이트나 보안이 강한 사이트
- 10초 이상: IP 차단 위험이 높은 사이트

재시도 기능

- 네트워크 오류 시 최대 3회 자동 재시도
 - 재시도 간격: 점진적 증가 (1초, 2초, 4초)
-

사이트별 크롤링 팁

네이버 쇼핑

- URL 예시: `https://shopping.naver.com/search/all?query=노트북`
- 권장 모드: selenium
- 크롤링 간격: 3-5초
- 수집 데이터: 상품명, 가격, 링크, 이미지

인스타그램

- URL 예시: `https://www.instagram.com/explore/tags/맛집/`
- 권장 모드: selenium (필수)
- 크롤링 간격: 5-10초
- 주의: 로그인 필요할 수 있음

부동산 사이트

- URL 예시: 직방, 다방, 네이버 부동산
- 권장 모드: selenium
- 크롤링 간격: 3-5초
- 수집 데이터: 매물 정보, 가격, 위치

일반 뉴스 사이트

- 권장 모드: requests
 - 크롤링 간격: 1-2초
 - 수집 데이터: 제목, 내용, 날짜, 링크
-

문제 해결

자주 묻는 질문

Q: 크롤링이 시작되지 않아요 A: - 인터넷 연결을 확인하세요 - URL이 올바른지 확인하세요 - 방화벽이나 백신 프로그램이 차단하고 있지 않은지 확인하세요

Q: 데이터가 수집되지 않아요 A: - selenium 모드로 변경해보세요 - 크롤링 간격을 늘려보세요 - 해당 사이트에서 크롤링을 차단하고 있을 수 있습니다

Q: 프로그램이 멈춰요 A: - 작업 관리자에서 프로그램을 종료하고 다시 시작하세요 - selenium 모드 사용 시 Chrome 브라우저가 자동으로 설치됩니다

Q: 엑셀 파일이 저장되지 않아요 A: - 저장 경로에 쓰기 권한이 있는지 확인하세요 - 같은 이름의 파일이 열려있지 않은지 확인하세요

오류 메시지별 해결법

"ModuleNotFoundError" - 프로그램을 다시 설치하세요 - 관리자 권한으로 실행해보세요

"TimeoutException" - 크롤링 간격을 늘려보세요 - 인터넷 연결 속도를 확인하세요

"Access Denied" - 해당 사이트에서 크롤링을 차단하고 있습니다 - VPN 사용을 고려해보세요

법적 고지사항

- 웹 크롤링 시 해당 사이트의 이용약관을 준수하세요
- 개인정보나 저작권이 있는 콘텐츠 수집 시 주의하세요
- 과도한 요청으로 서버에 부하를 주지 마세요
- 수집한 데이터의 사용에 대한 책임은 사용자에게 있습니다

SmartCrawler Pro v1.0.0

개발사: CrawlTech Solutions

문의: support@crawltech.com