

| 카카오톡 봇 강좌 | >

# [중급] [메봇R] API2에서 크롤링(파싱)을 해보자 (1)



마른얼음 1:1 채팅

2020.03.06. 11:59 조회 360

댓글 10 URL 복사

## 강좌할 주제

=> 음... Http 라이브러리?

## 강좌 내용

=> 오늘 나온(?)건지는 모르겠지만 따끈따끈한 API\_2에서 크롤링하는 법을 알아보시다.

일단 API\_2에 있는 크롤링 관련 라이브러리는 Http인데요, 이 객체의 메서드들을 한번 살펴봅시다.

```
Http.request(url | option, func);
Http.requestSync(url | option);
```

첫 번째 메서드 Http.request는 두 가지 사용법으로 나뉘는데요, 첫 번째 인자에 문자열로 링크를 주냐, json으로 옵션들을 주냐입니다. 첫 번째 인자를 링크로 준다면, 해당 사이트에 get 으로 크롤링하고, json으로 준다면 해당 옵션들로다가 request를 보냅니다. 이 json의 구조는..

```
{
  "url" : "링크", //크롤링할 사이트
  "timeout" : 3000, //타임아웃 시간 설정
  "method" : "GET", //GET or POST
  "headers" : { //헤더들
    "Accept" : "*/*",
    "User-Agent" : "Mozilla/5.0"
  }
}
```

대강 이렇게 되어있습니다. 헤더에는 User Agent와 Referer만 지원했었지만, 최근 업데이트로 모든 종류의 헤더를 지원한다고 하네요!

그리고 두 번째 인자에는 콜백함수가 들어갑니다. 즉 크롤링이 끝나면 두 번째 인자로 넘겨준 함수를 호출한다는 소리입니다. 이 콜백함수의 인자에는 exception, response, document가 들어갑니다. 이름에 드러나 있듯이 첫 번째 인자는 에러, 두 번째 인자는 http 응답, 세 번째 인자는 Jsoup으로 크롤링한 결과의 html이 전달됩니다.

이제 두 번째 메서드의 사용법을 알아보까요?

첫 번째 메서드와 약간은 비슷하지만, 이 requestSync라는 녀석은 동기적입니다. 인자에 url이나 json으로 옵션을 넘겨주면 Jsoup으로 크롤링한 결과의 html을 반환해줍니다. 첫 번째 메서드는 콜백함수를 수행하는 방식으로 작동하는 반면, 이 메서드는 html을 반환해준다는 차이가 있네요!

이제 이걸로 한강 수온을 가져오는 예제 소스를 짜 볼까요?

사이트는 wpws.kr 에서 가져오는걸로 하고, 크롬 개발자 도구를 이용하여 어디를 select할지 복사해웁시다. 그리고 이제 소스를 끄적이면 됩니다!

```
//Http.request 메서드 사용
Http.request("https://www.wpws.kr/hangang/", (e, r, d) => {
  if (e != null) {
    Log.e("한강 수온을 가져오지 못했습니다.\n" + e);
    return;
  }
  let temp = d.select("#temp").text();
  Log.d("현재 한강 수온 : " + temp);
});
```

//Http.requestSync 메서드 사용

```
try {
  let temp = Http.requestSync("https://www.wpws.kr/hangang/").select("#temp").text();
  Log.d("현재 한강 수온 : " + temp);
} catch (e) {
  Log.e("한강 수온을 가져오지 못했습니다.\n" + e);
}
```

개인적으로 Jsoup 라이브러리를 사용할 때보다 더 편리해진 것 같습니다..?

이상

2편은 안올라올수도

( 말머리 등록 해주세요! )



마른얼음님의 게시글 더보기 >

❤️ 좋아요 0 💬 댓글 10

🔗 공유 | 신고

댓글 등록순 최신순 ↻

댓글알림 ☐



delta

파일은 어떻게 올리나요

2020.03.06. 12:00 답글쓰기



마른얼음 작성자

흠터

2020.03.06. 12:02 답글쓰기



타카

따끈따끈한 고인물

2020.03.06. 12:02 답글쓰기



마른얼음 작성자

음?

2020.03.06. 12:04 답글쓰기



타카

마른얼음 마른얼음은 언제녹조 ㅋㅋ루ㅋㅋ

2020.03.06. 12:07 답글쓰기

삭제된 댓글입니다.



마른얼음 작성자

히히히히 | 히히히히 | 히히

2020.03.06. 12:19 답글쓰기



줄려

저번에 API2 뒤져보면서 써보고 노드느낌을 느꼈던건 안비밀

2020.03.06. 13:34 답글쓰기



마른얼음 작성자

ㅋㅋㅋㅋㅋㅋ

2020.03.06. 13:35 답글쓰기



Violet

질문 있습니다. 마른얼음은 왜 축척하지 않은건가요?

2020.03.06. 13:39 답글 쓰기



마른얼음 작성자

당연히 말라있으니까요(?)

2020.03.06. 13:39 답글 쓰기

Hibot

댓글을 남겨보세요



등록

글쓰기

답글

목록

▲ TOP

## '카카오톡 봇 강좌' 게시판 글

전체 [중급] 말머리 글

이 게시판 새글 구독하기 ☐

[초급] [메봇R] 따끈따끈한 API2 사용법을.araboza (2) [6]	마른얼음	2020.03.07.
[초급] [메봇R] 따끈따끈한 API2 사용법을.araboza (1) [5]	마른얼음	2020.03.06.
[중급] [메봇R] API2에서 크롤링(파싱)을 해보자 (1) [10]	마른얼음	2020.03.06.
카카오링크 커스텀 템플릿으로 더욱 알차게 쓰자! (KakaoLink 강좌 #3) 🤖 [21]	ApplePlus	2020.03.05.
카카오링크 커스텀 템플릿으로 더욱 알차게 쓰자! (KakaoLink 강좌 #2) 🤖 [25]	ApplePlus	2020.03.05.

1 2 3

전체보기

## 이 카페 인기글

eval 질문

Kiri  
♡0 💬5

안녕하세요

tomohong  
♡1 💬3

넌머ㅣㅣㅣㅣㅣㅣㅣㅣㅣㅣㅣㅣ  
ㅣㅣ  
[확습] 모니카는 여기에 잘 있음  
넌머ㅣㅣㅣㅣㅣㅣㅣㅣㅣㅣㅣㅣ  
[카카오톡 봇] 양산형 자동학습 봇 소스 공유



혹시 여기서 변수를 저장하는 부분이 어딘지 아시는분 계시나요..??

틱택토 (Player vs Bot)



틱택토 (카카오링크 적용)

마녀사냥 금지

AlphaDo  
♡3 💬8

공식 카카오톡 유허

Lunar  
♡0 💬9

덧셈/그룹챗 구분

시초는중국발  
♡0 💬8

