

[| 카카오톡 봇 강좌 | >](#)

## [응용] 메봇에서 스크립트가 로드된 상태의 웹페이지 파싱하기



산하 1:1 채팅

2020.06.18. 23:42 조회 149

댓글 7 URL 복사 ⋮

### 강좌할 주제

=> 자바스크립트가 로드된 후의 상태인 웹페이지 파싱하기

### 강좌 내용

=>

주의 : 이 강좌는 Node.js를 다룹니다.

다들 한 번쯤 웹페이지의 HTML 코드를 가져오면서, 분명 자신이 페이지에 접속해서 직접 눈으로 확인했던 값이 가져왔던 코드에는 눈을 씻고 찾아봐도 보이지 않는 경우가 있었을 겁니다. 그 이유중 하나가, 자바스크립트가 로드되지 않았기 때문입니다. 사이트가 완전히 로드된 후 자바스크립트 통신(?)을 이용하여 값을 주고받았던 것이죠.

이 강좌에서는 그 자바스크립트가 로드된 상태의 HTML 코드를 가져와보도록 하겠습니다.

1. 일단, <https://repl.it> 에 회원가입을 합니다. (온라인 IDE입니다.)
2. 회원가입 후에 Create a repl 에서 Node.js를 선택한 후 앱(?)을 생성해줍니다.
3. 생성된 앱 분명히 앱은 아닌거같은데 그냥 앱이라고 부르겠습니다 의 index.js 에 아래와 같은 코드를 입력해줍니다. (참고로, npm install 과 package.json의 수정 과정은 필요 없습니다.)

```
const express = require('express');
const requestSync = require('request-sync');
const url = require('url');
const JSDOM = require('jsdom').JSDOM;
const app = express();

app.get('/', (req, res) => {
  const URL = url.parse(req.url, true).query.url;
  let body = requestSync({
    method: 'GET',
    url: URL
  }).body;
  let dom = new JSDOM(body);
  let window = dom.window;
  let document = window.document;
  document.onload = function () {
    res.send(document.body.innerHTML);
    res.end();
  }
});

app.listen(3000, () => {
  console.log('server started');
});
```

4. 중앙 상단에 보이는 Start 버튼을 눌러서 서버를 시작해줍니다.
5. 서버가 만들어지면, 우측 상단에 서버의 주소가 나옵니다. 그 주소를 복사해주세요.
6. 아까 복사했던 주소를 사용해서 아래 코드를 완성한 후 자유롭게 사용하시면 됩니다.

```
function GETAfterJavaScriptLoaded (url) {  
  const ru = '생성된_서버_주소';  
  return org.jsoup.Jsoup.connect(ru + '?url=' + url).get();  
}
```

7. 끝

( 말머리 등록 해주세요! )



산하님의 게시글 더보기 >

❤️ 좋아요 3 💬 댓글 7

🔗 공유 | 신고

댓글 등록순 최신순 ↻

댓글알림 ☐



jomin398

그럼 npm 으로 node js 모듈을 그 앱에 설치하고.... 봇에선 그 앱을 파싱하면되는거죠?(정보는 어떻게넘기는지....).

2020.06.18. 23:51 답글 쓰기



도미 doami2005

그 앱(?) 에서 서버에 정보를 전송하고 봇에서 그 서버를 파싱하면 됩니다

2020.06.18. 23:52 답글 쓰기



jomin398

도미 doami2005 링크보내면 파일 보내게하려고 합니다.

이는 원래 컴에서 html 서버열어서 컴이 nodejs 모듈 돌린후 파일을 그서버에서 올려서

봇으로 그 다운로드 링크를 파싱해서 받는건데  
그게 컴퓨터를 대체할수있나요?

2020.06.18. 23:57 답글 쓰기



도미 doami2005

Termux 에서 돌려도 되ㄴ

2020.06.18. 23:51 답글 쓰기



Dark Tornado

자바스크립트 측에서 어디서 값을 가지고오는지 찾아서 뜯어오는 방법도 있는 거시예오.

2020.06.19. 00:26 답글 쓰기



비닐봉지

강의 해주실 닥토님 구합니다

2020.06.19. 06:08 답글 쓰기



bass9030

와 내가 옛날에 저거 할려고 구글에서 해엄치다 못구하고 포기했던 기억이 난다  
잘쓰겠습니다 :)

2020.06.19. 20:23 답글 쓰기

Hibot

댓글을 남겨보세요



등록

'카카오톡 봇 강좌' 게시판 글

전체    [응용] 말머리 글

이 게시판 새글 구독하기 ☐

[팁] JSFu*k if문(재업) <span>[6]</span>	1급 청정수 Moo nBot	2020.06.19.
[응용] 전역변수 응용 - 엽다운 게임 만들기 ☹️ <span>[15]</span>	도미 doami2005	2020.06.19.
[응용] 메봇에서 스크립트가 로드된 상태의 웹페이지 파싱하기 <span>[7]</span>	산하	2020.06.18.
[고급] Rabin-Karp Algorithm ☹️ <span>[4]</span>	윤동원	2020.06.18.
[고급] 문자열 탐색 알고리즘 (KMP) ☹️ <span>[5]</span>	윤동원	2020.06.18.

이 카페 인기글

로코가 뭐길래 사람들이 환장하는걸까?

쿨칠  
♡1   💬18

카봇커 채팅방을 그룹으로 만들면 오실분?

성빈  
♡1   💬52



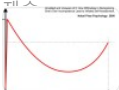
실시간 멜론 차트

안녕하십니까.

JSR  
♡0   💬8



웹툰



카톡봇 사람들은 꼭봐야됨



아키네이터

카툰 정지에 관해서

리스트  
♡0   💬15

아무래도 가입 나이 제한을 거는건 어떨까요...

공돌이  
♡0   💬33