


| 자유 게시판 | >

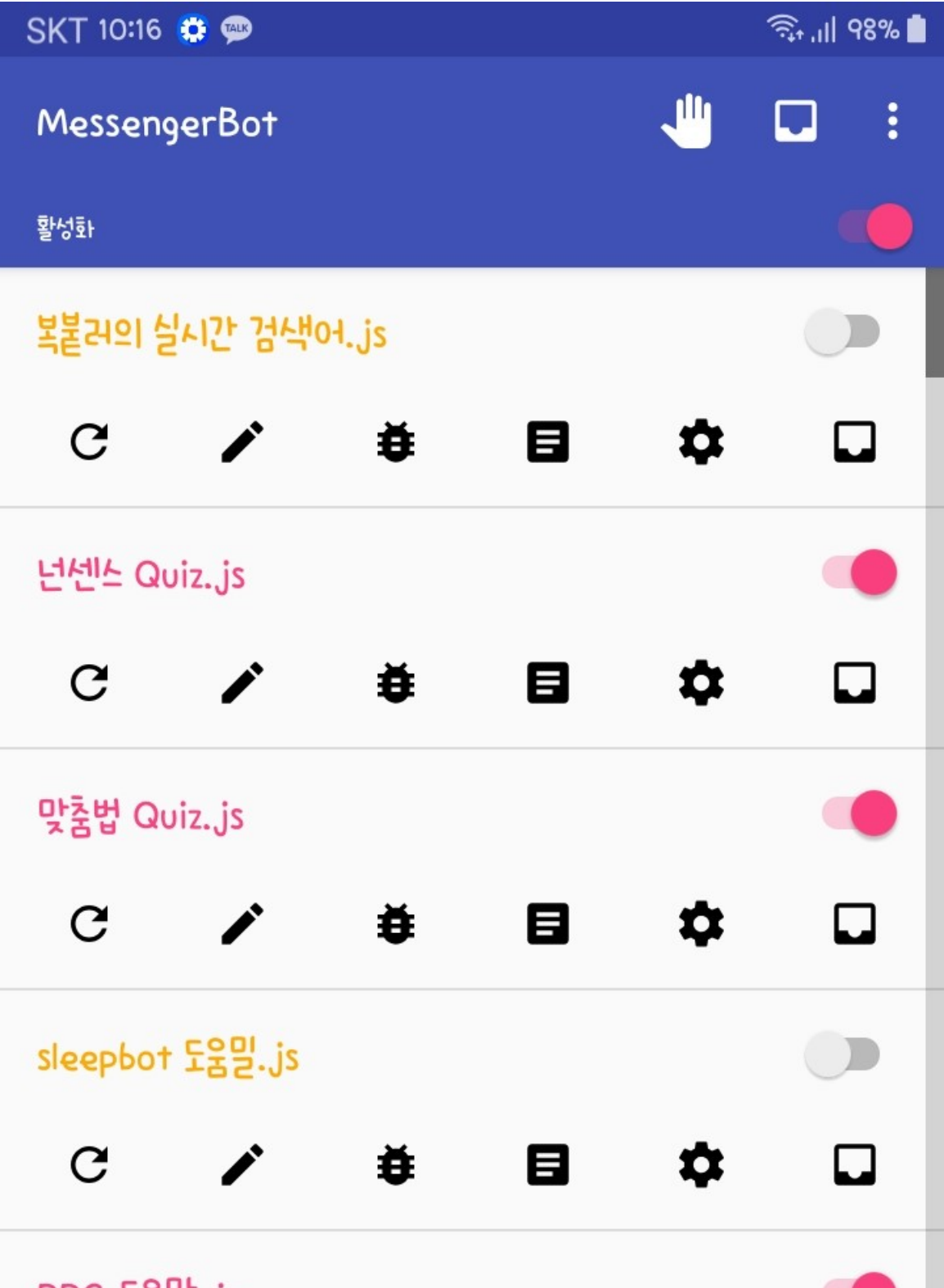
복불러 슬립이는 파싱을 배우는 중 2일차

 **슬립이** 챗봇 마스터 🍷 1:1 채팅

2019.09.08. 10:52 조회 180

💬 댓글 8 URL 복사 ⋮

안녕하세요! split만 배우려다가 replace, trim 까지 배우게 된 슬립이에요!
split, replace, trim 을 통틀어 정규식이라고 부르더군요!
html을 긁어왔을때 실시간 검색어가 아닌 인기 급상승으로 되어있기에 찾느라고 늦었습니다ㅠ
결국 완성을 했네요!



RPG 도움글.js



실시간 검색어 Naver.js



복불라의 실시간 검색어.js 라고 지정해주었습니다!

그럼 갓무위키를 도전할 때 처럼 일단 html을 긁어와줍시다!

```
function response(room, msg, sender, isGroupChat, replier, ImageDB, packageName, threadId){
var ranking = Utils.getWebText("https://naver.com")
if (msg == "킹갓네입버")
replier.reply(""+ ranking + "");
}
```

대충 긁어오면

SKT 10:18



97%



전체보기



```
<!doctype html>
<html lang="ko">
<head>
<meta charset="utf-8">
<meta name="Referrer" content="origin">
<meta http-equiv="Content-Script-Type" content="text/javascript">
<meta http-equiv="Content-Style-Type" content="text/css">
<meta http-equiv="X-UA-Compatible" content="IE=edge">
<meta name="viewport" content="width=1100">
<meta name="apple-mobile-web-app-title" content="NAVER">
<meta name="robots" content="index,nofollow">
<meta name="description" content="네이버 메인에서 다양한 정보와 유용한 콘텐츠를 만나 보세요">
<meta property="og:title" content="네이버">
<meta property="og:url" content="https://www.naver.com/">
<meta property="og:image" content="https://s.pstatic.net/static/www/mobile/edit/2016/0705/mobile_212852414260.png">
<meta property="og:description" content="네이버 메인에서 다양한 정보와 유용한 콘텐츠를 만나 보세요">
<meta name="twitter:card" content="summary">
<meta name="twitter:title" content="">
<meta name="twitter:url" content="https://www.naver.com/">
<meta name="twitter:image" content="https://s.pstatic.net/static/www/mobile/edit/2016/0705/mobile_212852414260.png">
<meta name="twitter:description" content="네이버 메인에서 다양한 정보와 유용한 콘텐츠를 만나 보세요">
<link rel="shortcut icon" type="image/x-icon" href="/favicon.ico?1">
<link rel="stylesheet" type="text/css" href="https://pm.pstatic.net/css/main_v190814.css">
<link rel="stylesheet" type="text/css" href="https://pm.pstatic.net/css/webfont_v170623.css">
<link rel="stylesheet" type="text/css" href="https://ssl.pstatic.net/ssstatic/search/pc/css/api_atcmp_190612.css">
<script type="text/javascript" src="https://pm.pstatic.net/js/c/nlog_v181107.js"></script>
<script type="text/javascript" src="https://ssl.pstatic.net/tveta/libs/assets/js/common/min/probe.min.js"></script>
<script type="text/javascript">
```

```

var nsc = "navertop.v3";
document.domain = "naver.com";
var jindoAll = "";
var iframeLazyLoad = false;
if (!!window.console) {window.console={};window.console["log"]=function() {}}
var isLogin = false;
function refreshLcs(etc) {etc = etc ? etc : {};if(document.cookie.indexOf("nrefreshx=1") != -1)
{etc["mrf"]="1";} else {etc["pan"]="bsn";}return etc;}

</script>
<title>NAVER</title>
<style>
#_nx_kbd .setkorhelp a { display:none; }
</style>
</head>
<body class="">
<!-- 스크립트 내비게이션 -->
<div class="u_skip">
<a href="#news_cast" onclick="document.getElementById('news_cast2').tabIndex =
-1;document.getElementById('news_cast2').focus();return false;">연습뉴스 바로가기</a>
<a href="#themecast" onclick="document.getElementById('themecast').tabIndex =
-1;document.getElementById('themecast').focus();return false;">주제별캐스트 바로가기</a>

```

히우엑 저번처럼 수많은 영어가 가득했습니다!
여기서 인기 급상승을 찾아야 하기에 내리고 내리면...

SKT 10:22



96%

← 전체보기



```

<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">14</span> <span class="ah_k">조찬형</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">15</span> <span class="ah_k">아스달 연대기</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">16</span> <span class="ah_k">하법 포이리에</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">17</span> <span class="ah_k">자동차범퍼속리달인</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">18</span> <span class="ah_k">동물농장</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">19</span> <span class="ah_k">부산라면달인</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">20</span> <span class="ah_k">맛있는 녀석들 뷔페</span> </a> </li>
</ul>
</div>
<div>
<span class="ah_ico_open"></span>
<div class="ah_list PM_CL_realtimeKeyword_list_base" aria-hidden="true">
<h3 class="ah_ltit">급상승 검색어</h3>
<a href="http://datalab.naver.com/keyword/realtimeList.naver?where=main" class="ah_ha" data-clk="lve.rankhistory"> <span class="ah_ico_datalab">DataLab.</span>급상승 트래킹<span class="ah_ico_hlink"></span> </a>
<div class="ah_tab">
<a href="#" role="tab" class="ah_tab_btn ah_tab_on" data-tab="1to10" data-clk="lve.tab1">1~10위</a>
<a href="#" role="tab" class="ah_tab_btn" data-tab="11to20" data-clk="lve.tab2">11~20위</a>
</div>
<h4 class="blind PM_CL_realtimeKeyword_sublist_heading">급상승 검색어 1~10위</h4>
<ul class="ah_l" data-list="1to10">
<li class="ah_item" data-order="1"> <a href="http://search.naver.com/

```



```

search.naver/?where=nexearch&query=%EB%A1%9C%EB%98%90875%ED%9A%8C%EB%8B%B9%
EC%B2%A8%EB%B2%88%ED%98%B8&sm=top_lve&ie=utf8" class="ah_a" data-ssl="https://
search.naver.com/
search.naver/?where=nexearch&query=%EB%A1%9C%EB%98%90875%ED%9A%8C%EB%8B%B9%
EC%B2%A8%EB%B2%88%ED%98%B8&sm=top_lve&ie=utf8" data-clk="lve.keyword") <span
class="ah_r">1</span> <span class="ah_k">로또875회당첨번호</span> </a> <a href="http://
datalab.naver.com/keyword/
realtimeDetail.naver?datetime=2019-09-08T10:18:00&query=%EB%A1%9C%EB%98%90875%ED
%9A%8C%EB%8B%B9%EC%B2%A8%EB%B2%88%ED%98%B8&where=main" class="ah_da" data-
clk="lve.kwdhistory"> <span class="blind">데이터랩 그래프 보기</span> <span class="ah_ico_datagraph"></
span> </a> </li>
<li class="ah_item" data-order="2"> <a href="http://search.naver.com/
search.naver/?where=nexearch&query=%EB%85%B8%EC%97%98&sm=top_lve&ie=utf8"
class="ah_a" data-ssl="https://search.naver.com/
search.naver/?where=nexearch&query=%EB%85%B8%EC%97%98&sm=top_lve&ie=utf8"
data-clk="lve.keyword"> <span class="ah_r">2</span> <span class="ah_k">노엘</span> </a> <a
href="http://datalab.naver.com/keyword/
realtimeDetail.naver?datetime=2019-09-08T10:18:00&query=%EB%85%B8%EC%97%98&w
here=main" class="ah_da" data-clk="lve.kwdhistory"> <span class="blind">데이터랩 그래프 보기</span>
<span class="ah_ico_datagraph"></span> </a> </li>
<li class="ah_item" data-order="3"> <a href="http://search.naver.com/
search.naver/?where=nexearch&query=%EC%9E%A5%EC%A0%9C%EC%9B%90&sm=top_lve&
ie=utf8" class="ah_a" data-ssl="https://search.naver.com/

```

여기부터 균요!

우리는 인기 급상승 검색어만 필요하기에

split으로 인기 급상승 위는 모두 없애줍시다!

.split("급상승 검색어 검색어")[1]

뒤에 [1] 은 급상승 검색어의 왼쪽을 모두 없애버리는 거더균요!

SKT 10:23   

  96% 

← 전체보기



```

</h3>
<div class="ah_roll_area PM_CL_realtimeKeyword_rolling">
<ul class="ah_l">
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">1</
span> <span class="ah_k">로또875회당첨번호</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">2</
span> <span class="ah_k">노엘</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">3</
span> <span class="ah_k">장제원</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">4</
span> <span class="ah_k">가라타 에리카</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">5</
span> <span class="ah_k">태이</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">6</
span> <span class="ah_k">강은일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">7</
span> <span class="ah_k">조찬형</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">8</
span> <span class="ah_k">오대환</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">9</
span> <span class="ah_k">코스트코 휴무일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">10</
span> <span class="ah_k">아이관</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">11</
span> <span class="ah_k">이마트 휴무일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">12</

```

```

span> <span class="ah_k">홈플러스 휴무일</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">13</
span> <span class="ah_k">롯데마트 휴무일</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">14</
span> <span class="ah_k">테이 매니저</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">15</
span> <span class="ah_k">아스달 연대기</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">16</
span> <span class="ah_k">하법 포이리에</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">17</
span> <span class="ah_k">자동차범퍼수리달인</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">18</
span> <span class="ah_k">동물농장</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">19</
span> <span class="ah_k">맛있는 녀석들 뷔페</span> </a> </li>
  <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">20</
span> <span class="ah_k">대형마트 휴무일</span> </a> </li>
</ul>
</div>
</div>
<span class="ah_ico_open"></span>
<div class="ah_list PM_CL_realtimeKeyword_list_base" aria-hidden="true">
  <h3 class="ah_tit">급상승 검색어</h3>
  <a href="http://datalab.naver.com/keyword/realtimeList.naver?where=main" class="ah_ha" data-
clk="lve.rankhistory"> <span class="ah_ico_datalab">DataLab.</span>급상승 트래킹<span
class="ah_ico_hlink"></span> </a>
  <div class="ah_tab">

```

결과는 이랬습니다! 아직까지는 순조롭게 진행이 되고 있군요!

그런데 우리는 급상승 검색어만 필요하지 그 밑에 링크나 도움말은 필요 없기에

.split("급상승 검색어")[0]

으로 급상승 검색어의 오른쪽 부분을 모두 없애줍시다!

결과는 이 정도가 되겠죠?

SKT 10:24



96%

← 전체보기



```

</h3>
<div class="ah_roll_area PM_CL_realtimeKeyword_rolling">
  <ul class="ah_l">
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">1</
span> <span class="ah_k">로또875회당첨번호</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">2</
span> <span class="ah_k">노엘</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">3</
span> <span class="ah_k">창제원</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">4</
span> <span class="ah_k">테이</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">5</
span> <span class="ah_k">카라타 에리카</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">6</
span> <span class="ah_k">조찬형</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">7</
span> <span class="ah_k">강은일</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">8</
span> <span class="ah_k">오대환</span> </a> </li>
    <li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">9</
span> <span class="ah_k">코스트코 휴무일</span> </a> </li>

```



```

<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">10</span> <span class="ah_k">아이린</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">11</span> <span class="ah_k">이마트 휴무일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">12</span> <span class="ah_k">롯데마트 휴무일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">13</span> <span class="ah_k">홈플러스 휴무일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">14</span> <span class="ah_k">테이 매니저</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">15</span> <span class="ah_k">아스달 연대기</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">16</span> <span class="ah_k">하법 포이리에</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">17</span> <span class="ah_k">자동차범퍼수리달인</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">18</span> <span class="ah_k">동물농장</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">19</span> <span class="ah_k">대형마트 휴무일</span> </a> </li>
<li class="ah_item"> <a href="#" class="ah_a" data-clk="lve.keyword"> <span class="ah_r">20</span> <span class="ah_k">맛있는 녀석들 뷔페</span> </a> </li>
</ul>
</div>
</div>
<span class="ah_ico_open"></span>
<div class="ah_list PM_CL_realtimeKeyword_list_base" aria-hidden="true">
<h3 class="ah_tit">

```

하지만 여러개의 태그 때문에 굉장히 더럽습니다

그렇다면 여기서 replace를 써야합니다!

replace 의 뜻은 바꾸어 채우다 인데

.replace(/(<([>]+)>)/g, "")

를 치면 여러개의 태그를 "" 안에 공백으로 바꾸어 채운다는 뜻이 되겠죠?

결과는 이렇습니다

SKT 10:26    

  95% 

← 전체보기



- 1 롯데875회당첨번호
- 2 노엘
- 3 장제원
- 4 테이
- 5 조찬형
- 6 카라타 에리카
- 7 강은일
- 8 코스트코 휴무일
- 9 오대환
- 10 아이린
- 11 이마트 휴무일
- 12 롯데마트 휴무일
- 13 홈플러스 휴무일

14 테이 매니저
15 아스달 연대기
16 하법 포이리에
17 자동차범퍼수리달인
18 동물농장
19 대형마트 휴무일
20 맛있는 녀석들 뷔페

Hm....

하지만 앞이나 맨 위쪽에 Tab을 친것같은 이상한 공백들이 되어있네요..

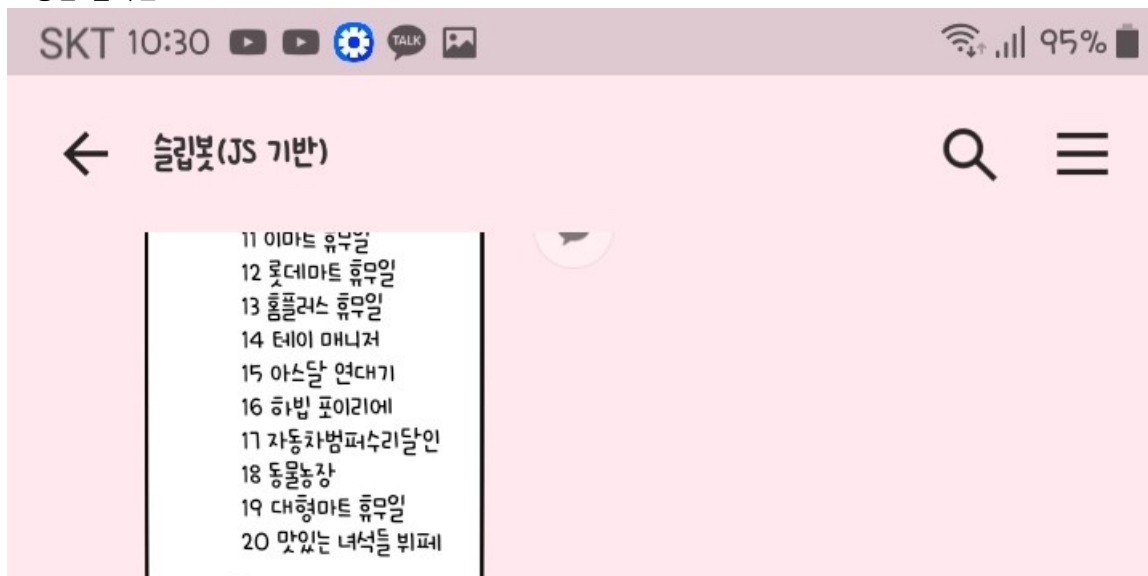
여기서 .trim 을 써 봅시다!

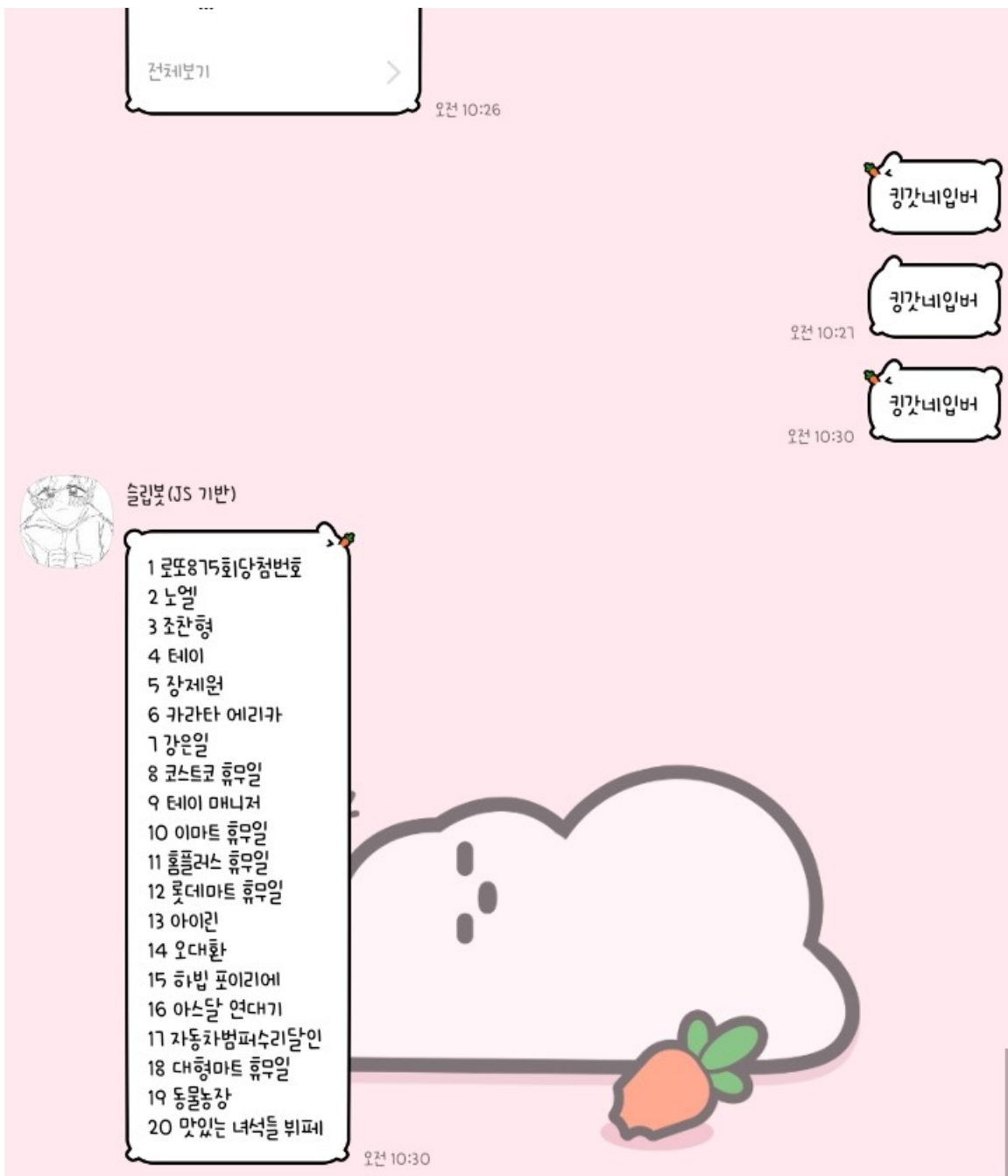
하지만 .trim 만 쓰면 줄바꿈이 되지 않기에

`.trim().replace(/[\\n\\s]{2,}/g,"\\n")`

하면 트림한 부분을 \\n 으로 줄바꿈을 한다는 뜻이 됩니다!

그렇담 결과는





정상적으로 작동 되는것을 볼 수 있습니다!

결국 슬립이도 완성 했습니다!

구글링과 카페의 글들을 모두 찾아서

복붙을 하지 않고 클리어 했습니다!

결국 결과 소스는 이렇게 되겠죠

```
function response(room, msg, sender, isGroupChat, replier, ImageDB, packageName, threadId){
var ranking = Utils.getWebText("https://naver.com").split("급상승 검색어 검색어")[1].split("급상승 검색어")
[0].replace(/(<([^\>]+)>)/g, "").trim().replace(/[\WnWs]{2,}/g, "\Wn")
if (msg == "킹갓네입버")
replier.reply(""+ ranking + "");
}
```

긴글 봐주셔서 감사합니다! 내일은 다른 사이트를 파싱해볼게요!

^~^



슬립이님의 게시글 더보기 >

❤ 좋아요 1 💬 댓글 8

🔗 공유 | 신고

댓글 등록순 최신순 ↻

댓글알림 ☐



사로로

곳곳 다음번엔 좀 더 어려운 사전 파싱해보시죠. 근데 파싱 소스는 if문 안에 넣으세요 저리 짜면 메시지 올때마다 파싱 되어서 데이터가 많이 소모됩니다

2019.09.08. 11:07 답글쓰기



슬립이 작성자

시도 해 보았지만..
저는 불가능 한 것 같더군요..
다른 분들 거 보니까 msg.sub어짜구..
하다보니 시도 해 보았더니 어디서 부터 시작하는지조차 알수 없어서 불가능 할 것 같네요ㅠㅠ

2019.09.09. 21:35 답글쓰기



MatSoGeum

정규식은 regex 아닌가ㅇ..

2019.09.08. 11:00 답글쓰기



사로로

regex 맞죠...

2019.09.08. 11:53 답글쓰기



슬립이 작성자

사로로 허형 구글에서 정규식이랬는데.. 쩌..

2019.09.08. 12:14 답글쓰기



초보인 사칭아닌 성빈

초보같은소리하고있네

2019.11.10. 10:46 답글쓰기



초보인 사칭아닌 성빈

내가더 초보인데(머쓱)

2019.11.10. 10:47 답글쓰기



하루

우오 ㅈ아아아아아아 ㅈ앙아 하루 왔다감

2020.04.02. 16:19 답글쓰기

Hibot

댓글을 남겨보세요



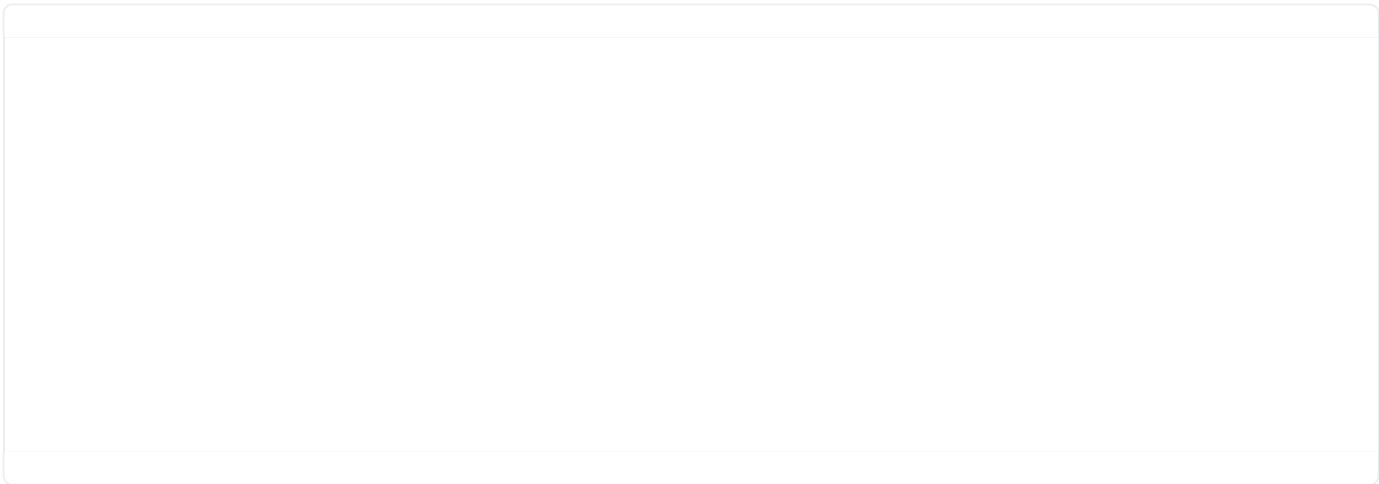
등록

✎ 글쓰기

답글

목록

▲ TOP



'자유 게시판' 게시판 글

이 게시판 새글 구독하기 ☐

웹사이트 삽질 현황 🙄 [7]	Dark Tornado	2019.09.08.
갑자기 궁금해진것 [3]	쿨칠	2019.09.08.
복불러 슬립이는 파싱을 배우는 중 2일차 🙄 [8]	슬립이	2019.09.08.
뭐만들어야 할까요 [5]	runkan	2019.09.08.
안자시는분 짜리질르엇	초보자	2019.09.08.

1 2 3

전체보기

이 카페 인기글

카페에 RPG가 많이 보이는 것 같군요

AlphaDo
♡1 💬3

중수가 초보에게 하는 강좌 [2]

하하하하
♡1 💬11



자동응답봇으로 RPG봇 만들었습니다



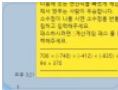
태양, 달 정보 구현 완료

Dark Tornado
♡0 💬10
하하하하
♡0 💬5



틱택토 (Player vs Player)

1 2 3 4



아이스봇 계산문제

error(재팅수(Left
Error: "FS" is not
#14) at
#14 (respons

채팅순위봇이 갑자기 안되네요..ㅠ

카카오톡 단순 자동응답봇 파일
도루묵
♡0 💬2