

SCENARIO GUIDE  
INTERMEDIATE LEVEL

: SCENARIO NAME

استخراج محتوا از وبسایت [theguardian.com](http://theguardian.com)

نسخه ی : ۱,۰,۱

تاریخ : ۱۳۹۶/۱۲/۰۳

## به نام خدا

### هدف

هدف این سناریو استخراج اخبار از وبسایت خبری **theguardian.com** و ایجاد تغییرات جزئی بر روی آن می باشد .

### شرح کلی سناریو

ژلیوگراف ابتدا به صفحه ی **RSS** مربوط به وبسایت **theguardian.com** مراجعه می کند و بخش هایی از خبر را که شامل عنوان ، خلاصه ، متن و منبع می شود ، توسط سامانه برای اولین بار انتخاب و استخراج می کند سپس درون یک رجیستر مجزا ذخیره می نماید .

### پیاده سازی سناریو

برای پیاده سازی این سناریو باید اقدامات زیر انجام گیرد :

- در هر مرحله ، یک خبر از قسمت **RSS** وبسایت انتخاب می شود که از ویژگی **Unique** برخوردار است و در دفعات قبلی اجرای سناریو استفاده نشده باشد .
- در صفحه ی خبر مورد نظر به ترتیب کادر مربوط به عنوان ، لید (خلاصه) ، متن و برچسب های (**Tag**) خبر یکی پس از دیگری انتخاب شده و سپس محتوای آن ها درون یک سری رجیستر ذخیره می شود .
- لینک تمامی تصاویر ابتدا استخراج شده سپس دائلود صورت می گیرد . با اینکار می توانیم پس از مدتی تمامی عکس ها را در سایت مقصد بارگذاری نماییم .
- در ابتدای متن خبر عبارت "به گزارش ژلیوگراف" اضافه می شود .

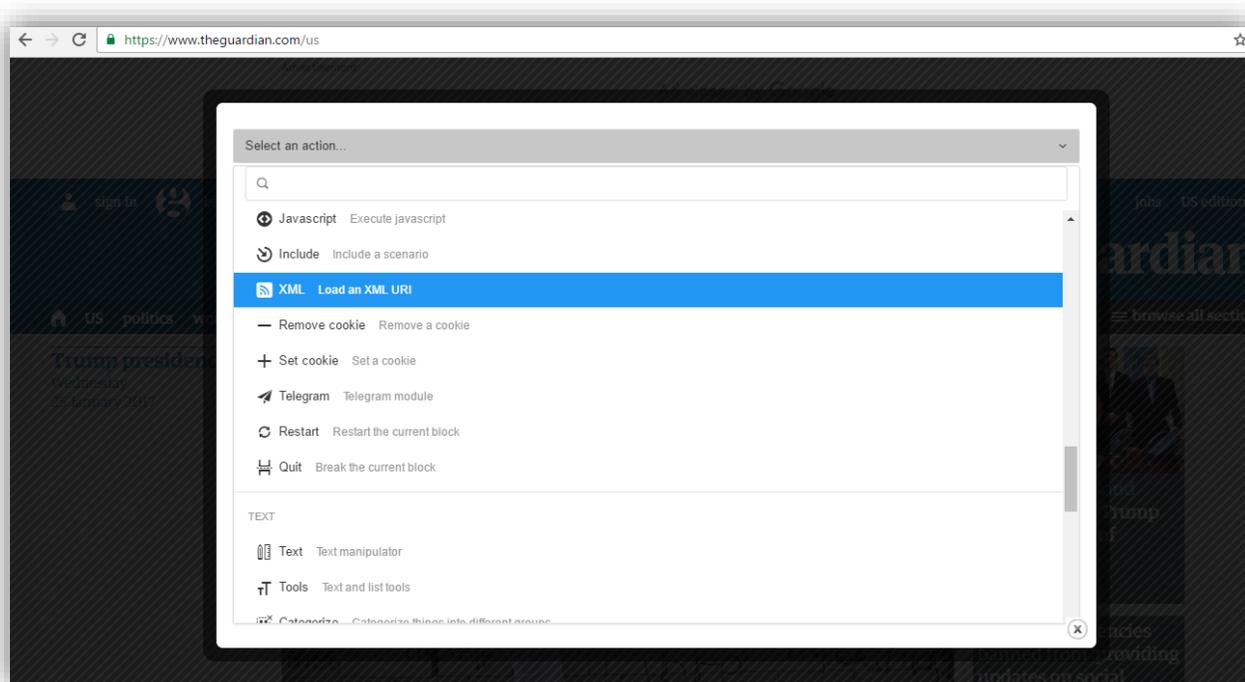
اگر کاربر بخواهد قسمت های مختلف یک خبر که به ترتیب شامل عنوان ، خلاصه ، متن ، بخش و برچسب ها است ، انتخاب کند و درون رجیستر های جداگانه ذخیره کند ، ابتدا باید صفحه ی خبر مورد نظر را باز کند . برای شبیه سازی این عمل توسط ژلیوگراف نیز لازم است خبرهای سایت به ترتیب باز شود و سپس قسمت های مورد نظر انتخاب شود و در ادامه هر کدام درون یک رجیستر مجزا قرار گیرد . به این منظور وبسایت خبری **theguardian.com** را می گشاییم .

## پلان صفر

در پلان **XML** ، آدرس **RSS** وبسایت خبری **theguardian.com** را قرار می دهیم و برای سامانه تعیین می کنیم تا با مراجعه به آدرس **RSS** وبسایت خبری **theguardian.com** خبرها را جهت استخراج محتوا به صورت منحصر به فرد و تصادفی انتخاب کند .

## گام نخست

پس از بارگذاری وبسایت خبری **theguardian.com** ، با فشردن کلید های **Alt+J** از کادر کشویی **Select an action** پلان **XML** را انتخاب می کنیم .



## گام دوم

آدرس [theguardian.com/rss](https://www.theguardian.com/rss) را که مربوط به صفحه ی RSS وبسایت [theguardian.com](https://www.theguardian.com) می باشد در فیلد XML Url وارد می کنیم .

به منظور این که اخبار منحصر به فرد و تصادفی انتخاب شوند گزینه های **Random** و **Unique** را فعال می نماییم و در انتها بر روی دکمه ی **Confirm** کلیک می کنیم .

The screenshot shows a mobile application interface titled "XML". It features a form with the following fields and controls:

- XML Uri:** A text input field containing the URL `https://www.theguardian.com/rss`.
- Action:** A dropdown menu set to "Open a link".
- Random:** A toggle switch that is turned on (green).
- Unique:** A toggle switch that is turned on (green).
- No feed strategy:** A dropdown menu set to "Success and end".
- Unique key (Advanced):** An empty text input field.
- Uniq similarity (Advanced):** A range input field showing "0 - 100, Default: Two texts would be considered identical with 75% of similarity".
- Containing:** An empty text input field.
- Plan:** A range input field showing "0".

At the bottom of the form, there are three buttons: "Refresh" (blue), "Back" (orange), and "Confirm" (green).

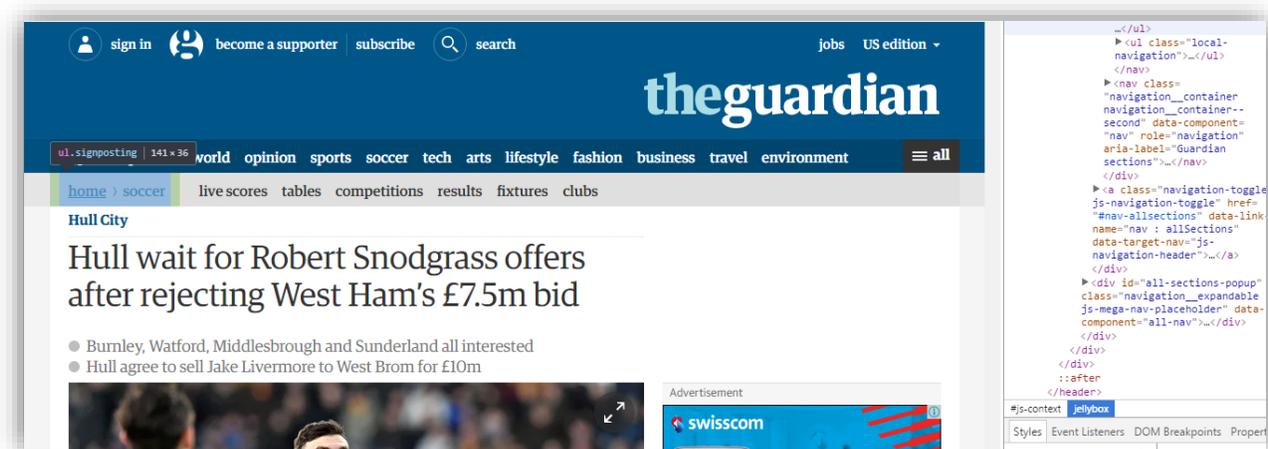
جهت استخراج محتوا وارد صفحه ی مربوط به یکی از اخبار می شویم .

## پلان یک

در نظر داریم تا سامانه با استفاده از پلان **Select** بخش(موضوع) مربوط به هر خبر را انتخاب کند و در پلان بعد آن را درون یک رجیستر ذخیره نماید .

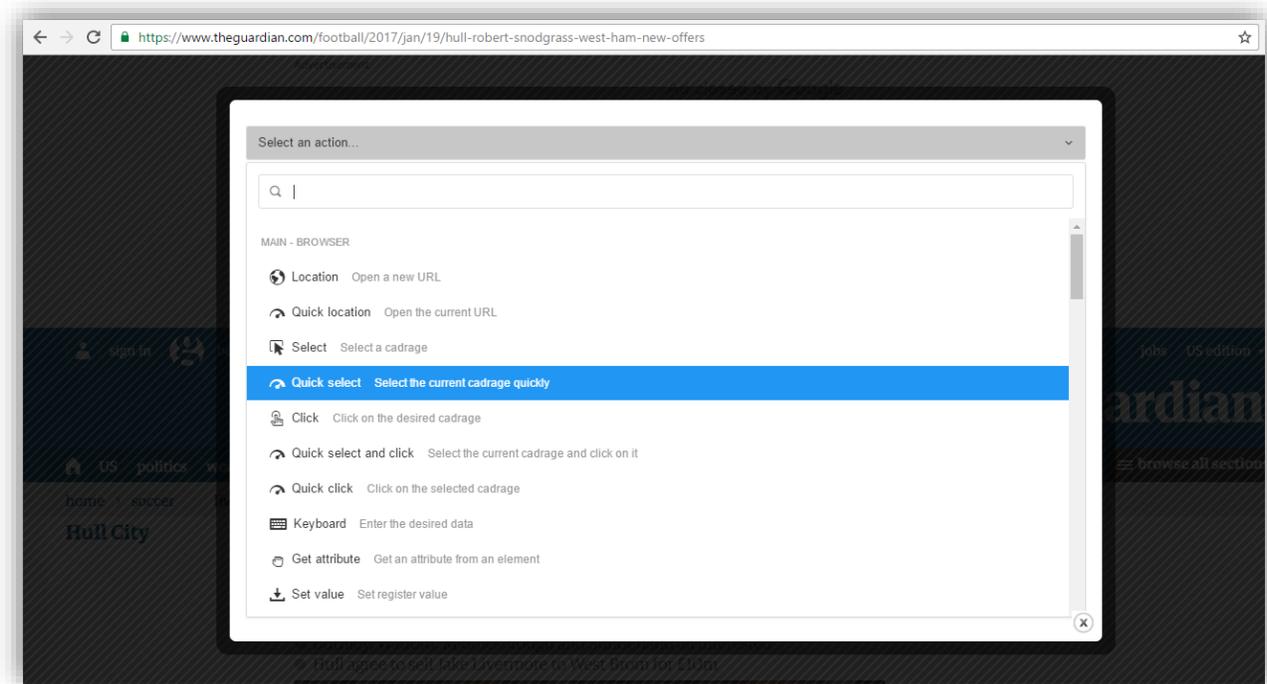
## کام نخست

با خارج شدن از سامانه ، ابتدا یک خبر را انتخاب کرده و بر روی عنوان آن کلیک کرده تا صفحه ی مربوط به آن باز شود ، سپس با فشردن کلیدهای **Ctrl+Shift+C** اشاره گر ماوس را بر روی عبارات مربوط به بخش خبر قرار می دهیم و یک بار کلیک می کنیم تا صفحه ی سامانه گشوده شود .



## گام دوم

از کادر کشویی **Select an action** گزینه ی **Quick select** را انتخاب می کنیم .

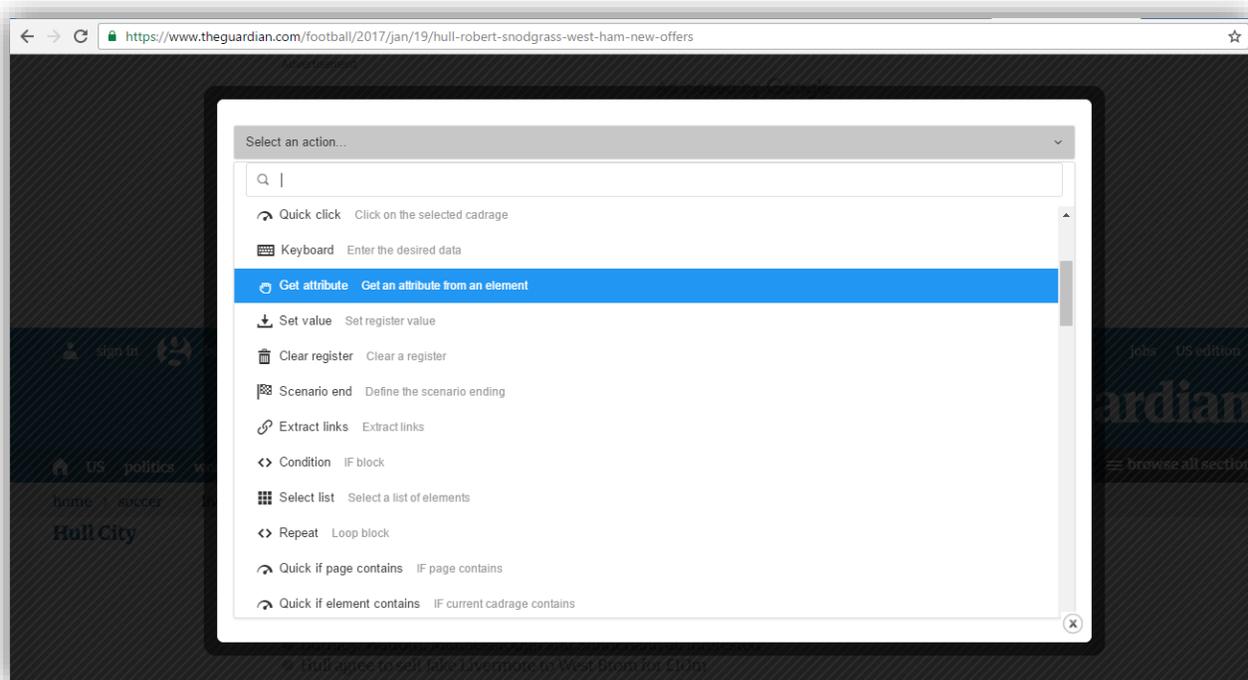


## پلان دو

در این قسمت لازم است تا ژلیوگراف به کمک پلان **Get attribute** بخش خبر انتخابی را درون یک رجیستر ذخیره کند .

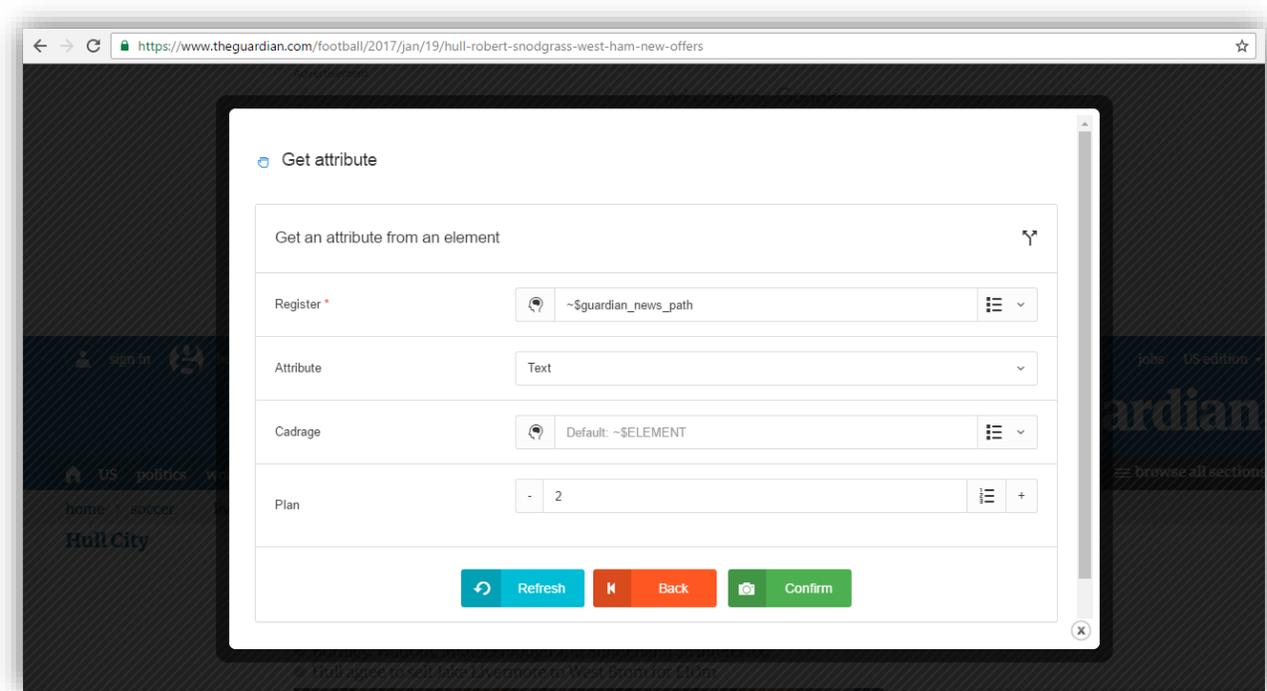
## کام نخست

بدون نیاز به خارج شدن از صفحه ی سامانه ، از کادر کشویی **Select an action** پلان **Get attribute** را انتخاب می کنیم.



## گام دوم

در فیلد **Register** نام رجیستر **~\$guardian\_news\_path** را قرار می دهیم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .



## پلان سه

می خواهیم ژلیوگراف با استفاده از پلان **Select** ، عنوان مربوط به خبر را انتخاب کند و در پلان بعد آن را درون یک رجیستر ذخیره نماید .

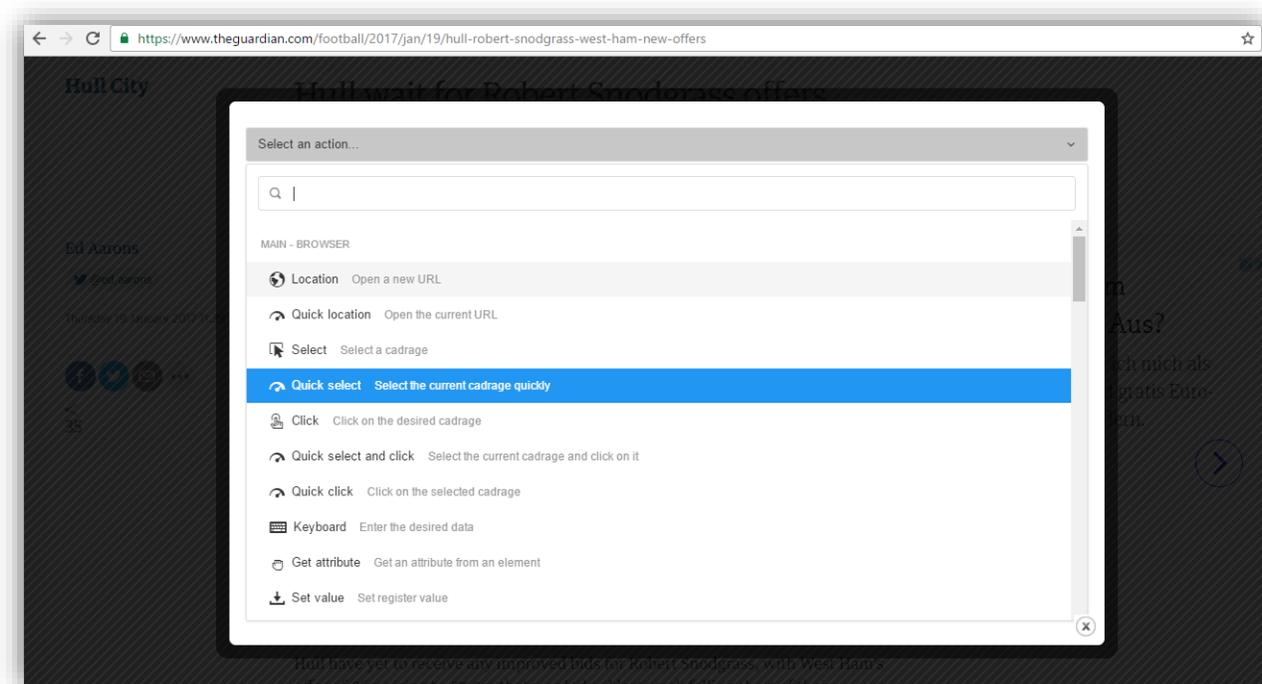
## گام نخست

با خارج شدن از سامانه ، پس از فشردن کلیدهای **Ctrl+Shift+C** | شماره گر ماوس را بر روی عبارات مربوط به عنوان خبر قرار می دهیم و یک بار کلیک می کنیم تا صفحه ی سامانه گشوده شود .

The screenshot shows a web browser displaying a news article from The Guardian. The article title is "Hull wait for Robert Snodgrass offers after rejecting West Ham's £7.5m bid". Below the title, there are two bullet points: "Burnley, Watford, Middlesbrough and Sunderland all interested" and "Hull agree to sell Jake Livermore to West Brom for £10m". The main image shows Robert Snodgrass in a Hull City kit kicking a ball. Below the image, there is a caption: "Robert Snodgrass refused to sign a new three-year contract at Hull last month. Photograph: Greig Cowie/BPI/REX/Shutterstock". The browser's developer tools are open on the right side, showing the HTML structure of the page. The selected element is an `h1` tag with the text "Hull wait for Robert Snodgrass offers after rejecting West Ham's £7.5m bid". The developer tools also show an advertisement for "Droht dem Euro das Aus?" from "gruener-fisher.de".

## گام دوم

از کادر کشویی **Select an action** گزینه **Quick select** را انتخاب می کنیم .

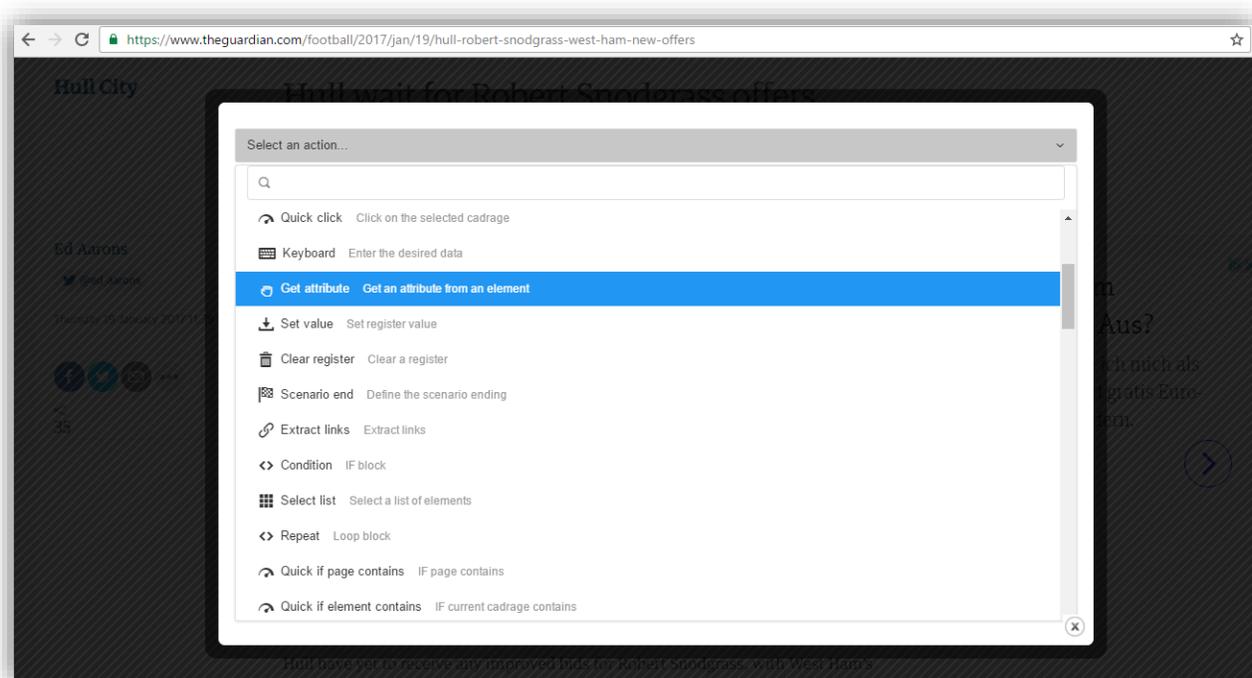


## پلان چهار

در این قسمت لازم است تا ژلیوگراف به کمک پلان **Get attribute** عنوان خبر را درون یک رجیستر ذخیره کند .

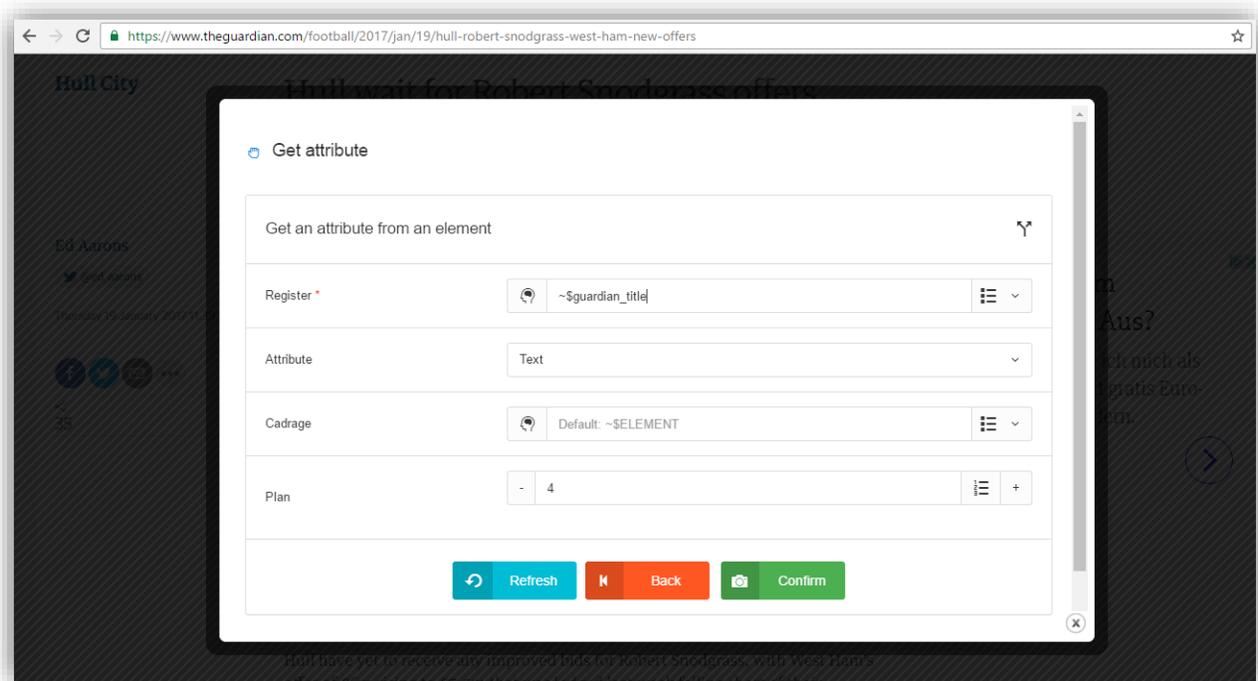
## کام نخست

بدون نیاز به خارج شدن از صفحه ی سامانه ، از کادر کشویی **Select an action** پلان **Get attribute** را انتخاب می کنیم.



## گام دوم

در فیلد **Register** نام رجیستر **~\$guardian\_title** را قرار می دهیم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .



## پلان پنچ

در نظر داریم تا ژلیوگراف با استفاده از پلان **Select** خلاصه ی خبر را انتخاب کند و در پلان بعد آن را درون یک رجیستر ذخیره نماید.

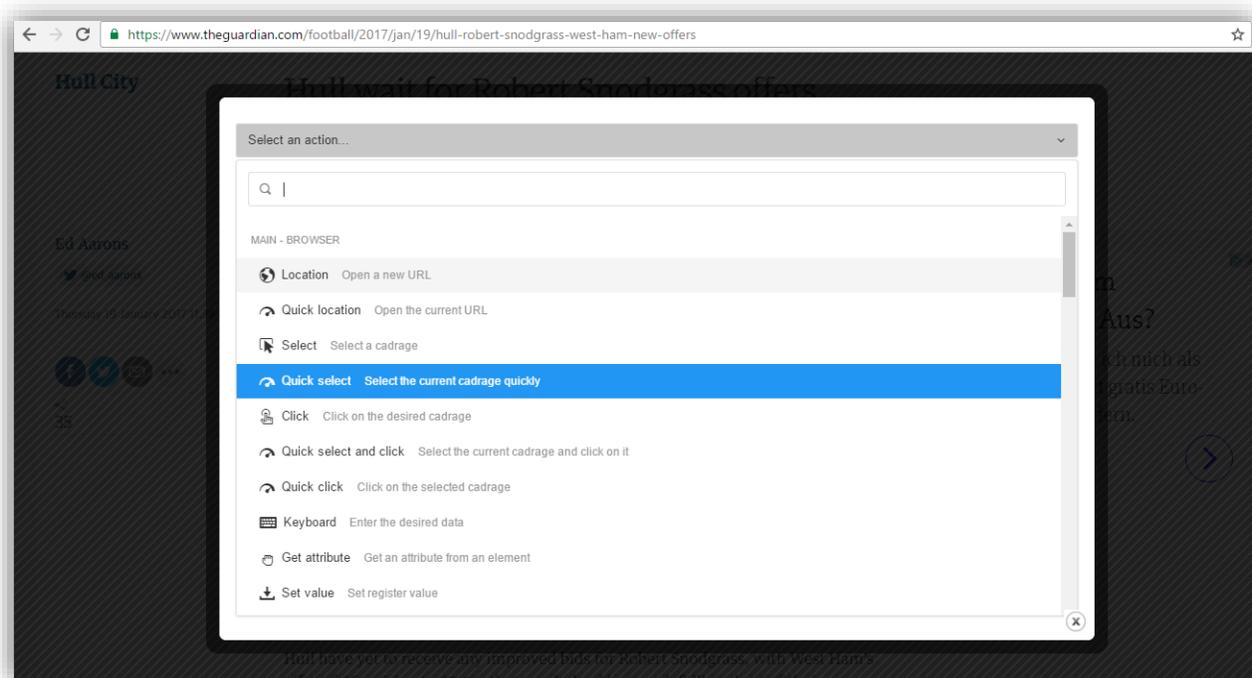
## گام نخست

با خارج شدن از سامانه ، پس از فشردن کلیدهای **Ctrl+Shift+C** | شماره گر ماوس را بر روی عبارات مربوط به خلاصه ی خبر قرار می دهیم و یک بار کلیک می کنیم تا صفحه ی سامانه گشوده شود .



## گام دوم

از کادر کشویی **Select an action** گزینه **Quick select** را انتخاب می کنیم .



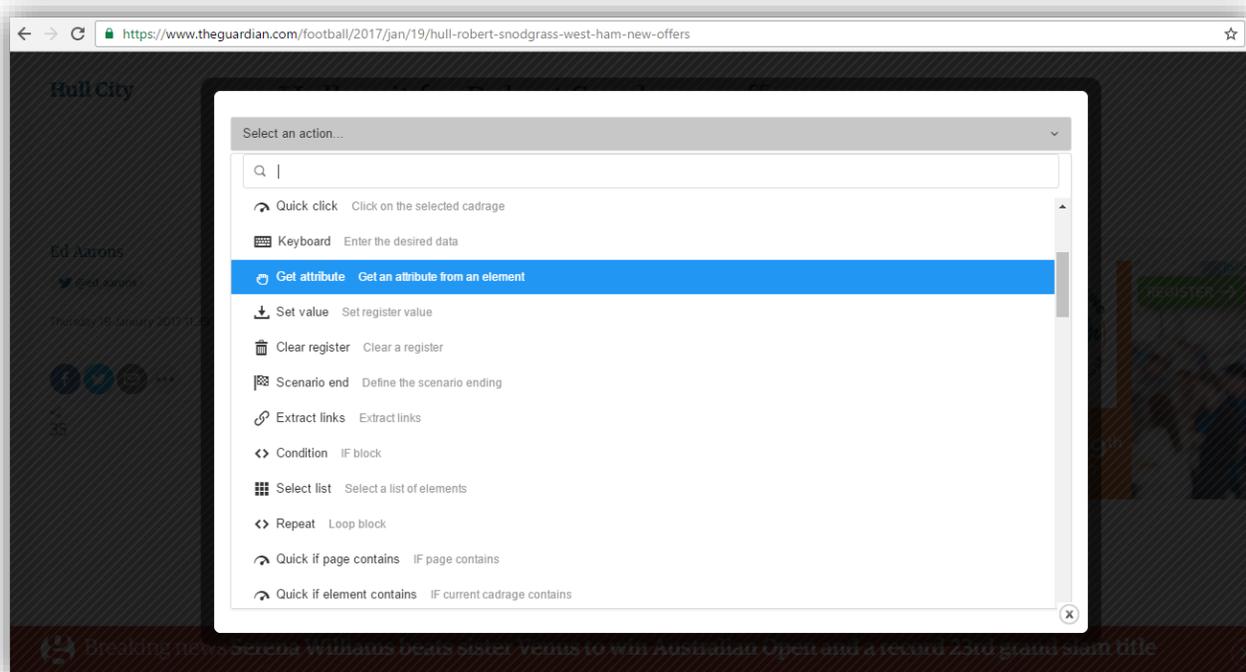
از آنجایی که ممکن است در برخی خبرها خلاصه وجود نداشته باشد ، به منظور جلوگیری از ایجاد خطا هنگام انتخاب خلاصه ی خبر می خواهیم تاکید کنیم تا تنها کادرهایی انتخاب شوند که کلاس آن ها برابر با کلاس خلاصه ی خبر باشد . چنانچه پس از فشردن کلید های **Ctrl+Shift+C** اشاره گر ماوس را بر روی خلاصه ی خبر قرار دهیم خواهیم دید که خلاصه ی خبر دارای کلاس **content\_\_standfirst** است . در این صورت چنانچه خلاصه ی خبر نداشته باشیم رجیستر **~\$guardian\_lead** بدون مقدار باقی می ماند .

## پلان ششی

در این قسمت لازم است تا ژلیوگراف به کمک پلان **Get attribute** عنوان خبر را درون یک رجیستر ذخیره کند .

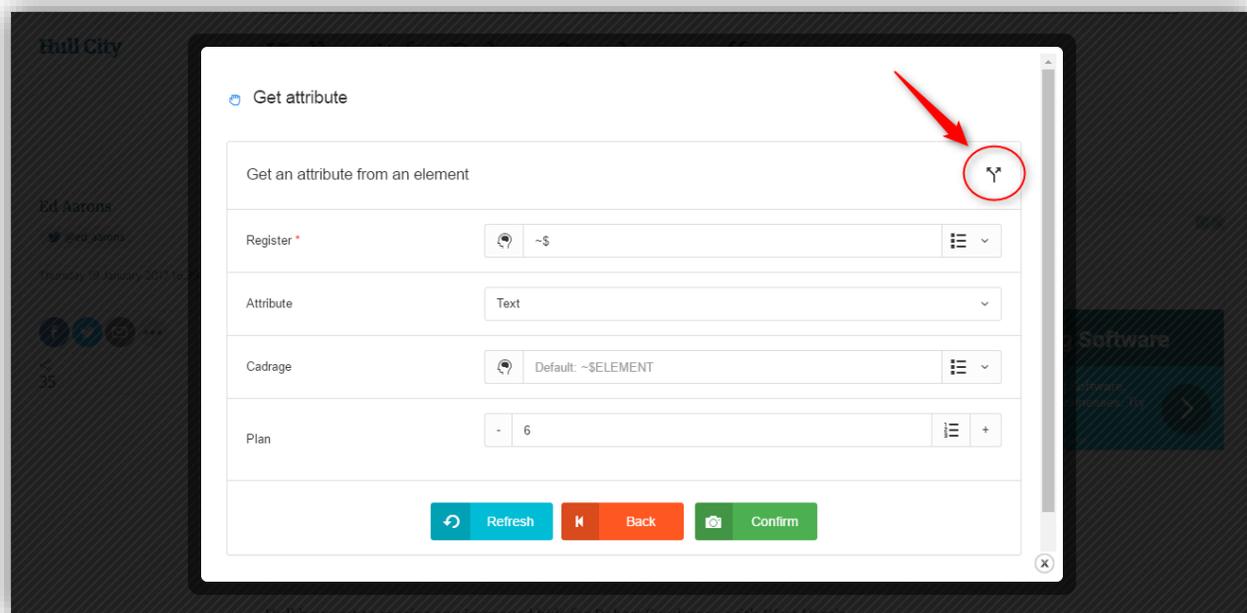
## کام نخست

بدون نیاز به خارج شدن از صفحه ی سامانه ، از کادر کشویی **Select an action** پلان **Get attribute** را انتخاب می کنیم.



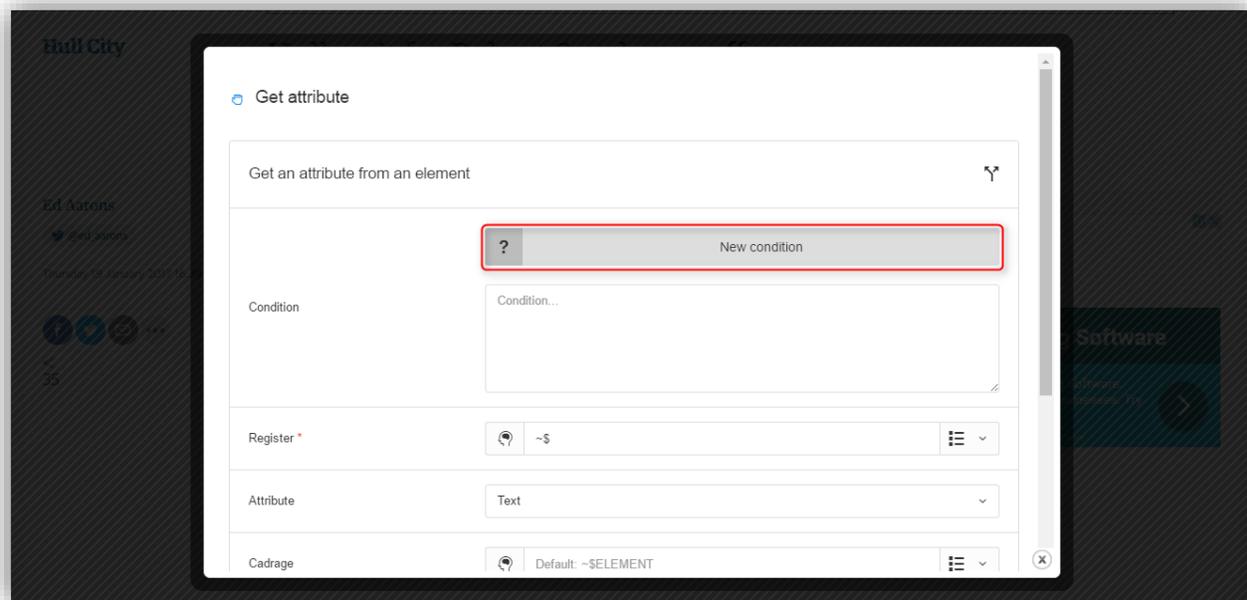
## گام دوم

بر روی علامت شرط یک بار کلیک می کنیم .



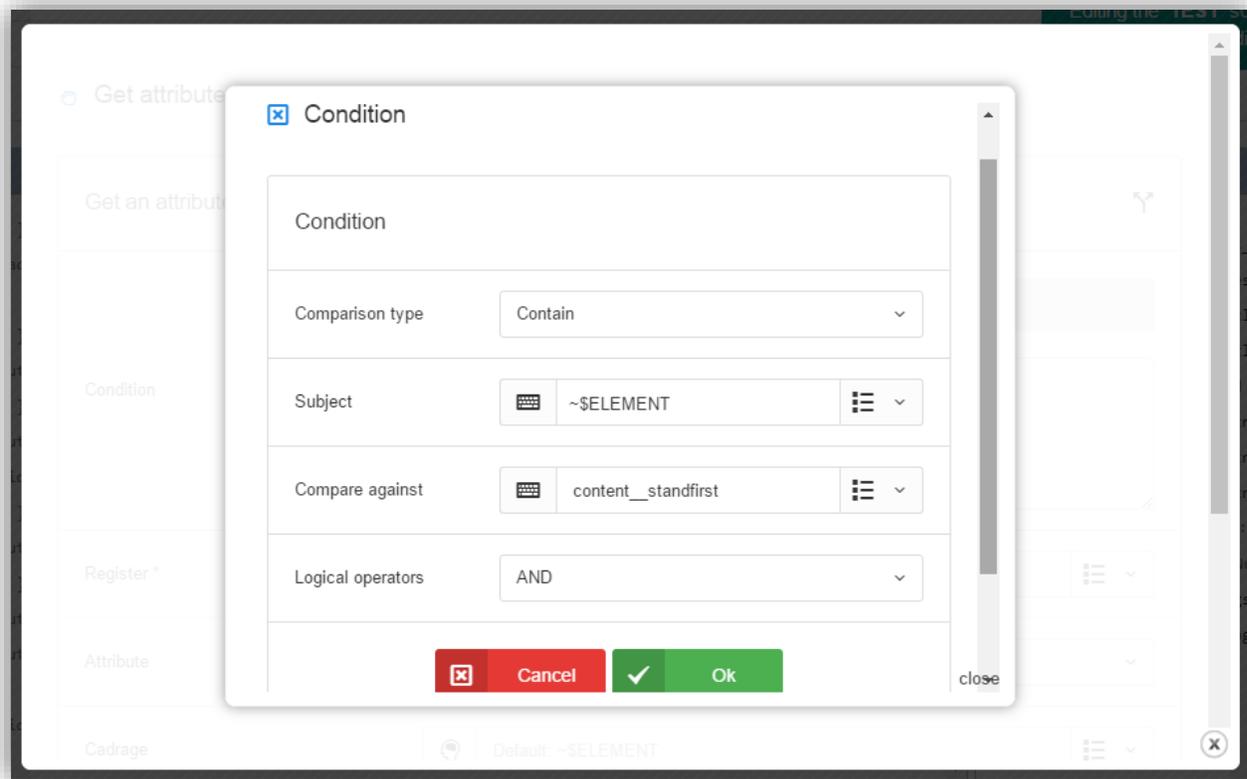
## گام سوم

کادر **New condition** را انتخاب می کنیم .



## گام چهارم

از کادر کشویی **Subject** گزینه **~\$ELEMENT** را انتخاب می کنیم و عبارت **content\_\_standfirst** را در کادر **Compare against** قرار می دهیم و در انتها بر روی گزینه **OK** کلیک می کنیم .



## گام پنجم

در فیلد **Register** نام رجیستر **~\$guardian\_lead** را قرار می دهیم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .

The screenshot shows the 'Get attribute' configuration screen. It includes the following fields and controls:

- Get an attribute from an element**: Title of the configuration.
- Condition**: A text area containing the expression `contain(~$SOURCE, 'content__standfirst') and`. Above the text area is a 'New condition' button with a question mark icon.
- Register \***: A dropdown menu with the selected value `~$guardian_lead`.
- Attribute**: A dropdown menu with the selected value `Text`.
- Cadrage**: A dropdown menu with the selected value `Default: ~$ELEMENT`.
- Plan**: A numeric input field with the value `6`.

At the bottom of the interface, there are four buttons: **Refresh** (circular arrow icon), **Back** (left arrow icon), **Inject** (plus icon), and **Confirm** (camera icon).

## پلان هفت

در نظر داریم تا ژلیوگراف با استفاده از پلان **Select** متن خبر را به همراه کدهای **HTML** انتخاب کند و در پلان بعد آن را درون یک رجیستر ذخیره نماید .

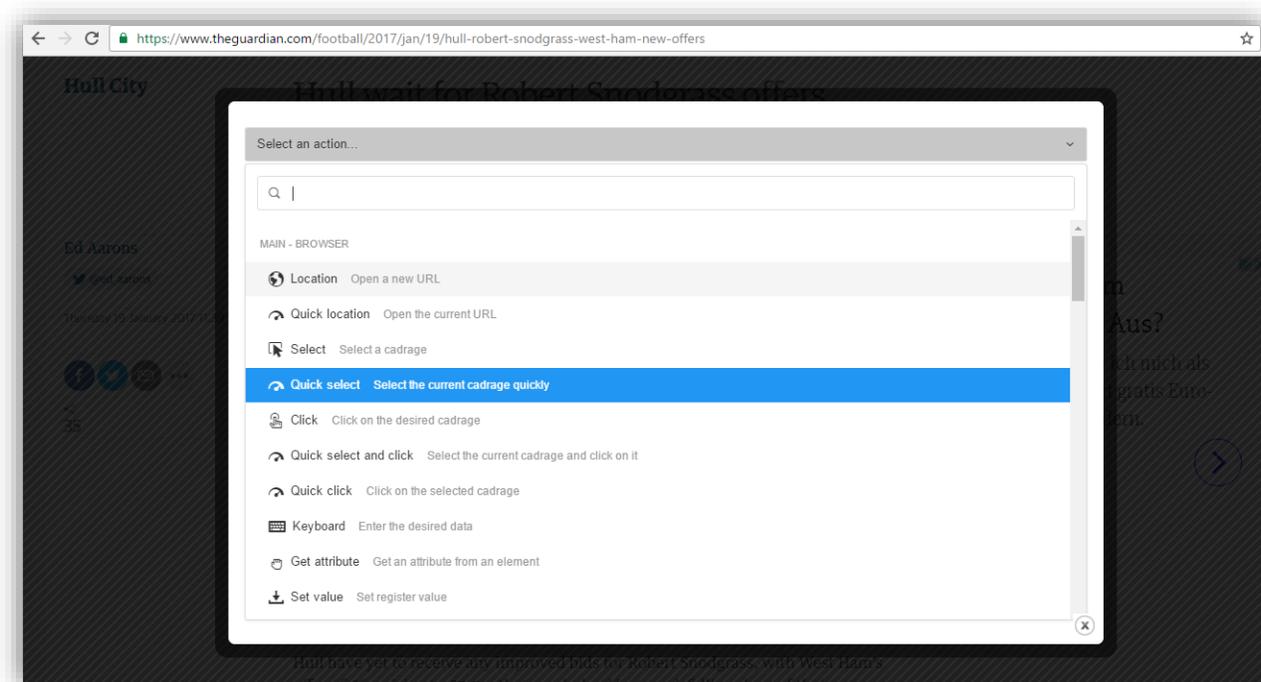
## گام نخست

با خارج شدن از سامانه ، پس از فشردن کلیدهای **Ctrl+Shift+C** اشاره گر ماوس را بر روی متن خبر قرار می دهیم و یک بار کلیک می کنیم تا صفحه ی سامانه گشوده شود .

The screenshot shows a web browser displaying a news article. The article title is "Hull agree to sell Jake Livermore to West Brom for £10m". The author is Ed Aarons. The article text discusses Robert Snodgrass's contract situation at Hull City, mentioning offers from West Ham and interest from other Premier League clubs like Burnley, Watford, Middlesbrough, and Sunderland. It also mentions Crystal Palace's interest in Snodgrass and Yohan Cabaye. A "Most popular in US" section is visible at the bottom of the article, featuring a headline about Berlin mayor Donald Trump. On the right side of the browser, the developer console is open, showing the "Elements" panel with a tree view of the page's DOM. The selected element is a paragraph of text, and its HTML code is displayed in the right pane. The code includes various classes and attributes, such as "content\_article-body" and "articleBody".

## گام دوم

از کادر کشویی **Select an action** گزینه **Quick select** را انتخاب می کنیم .

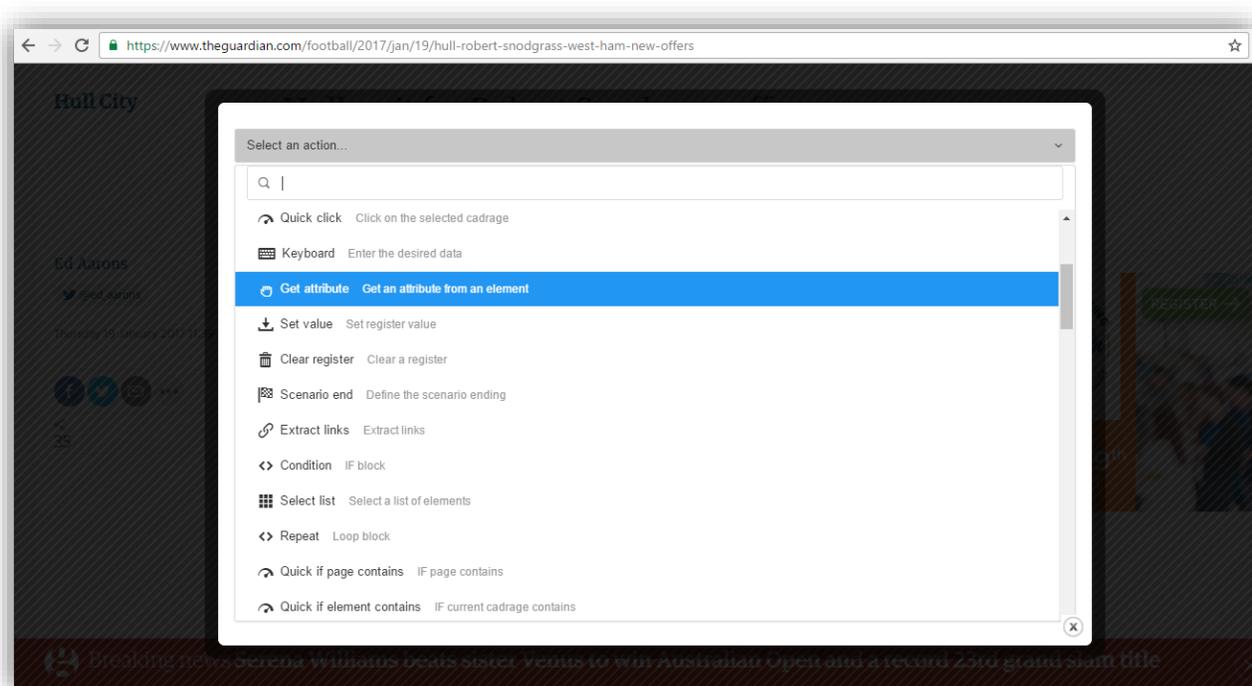


## پلان هشت

در این قسمت می خواهیم ژلیوگراف به کمک پلان **Get attribute** متن خبر را درون یک رجیستر ذخیره کند .

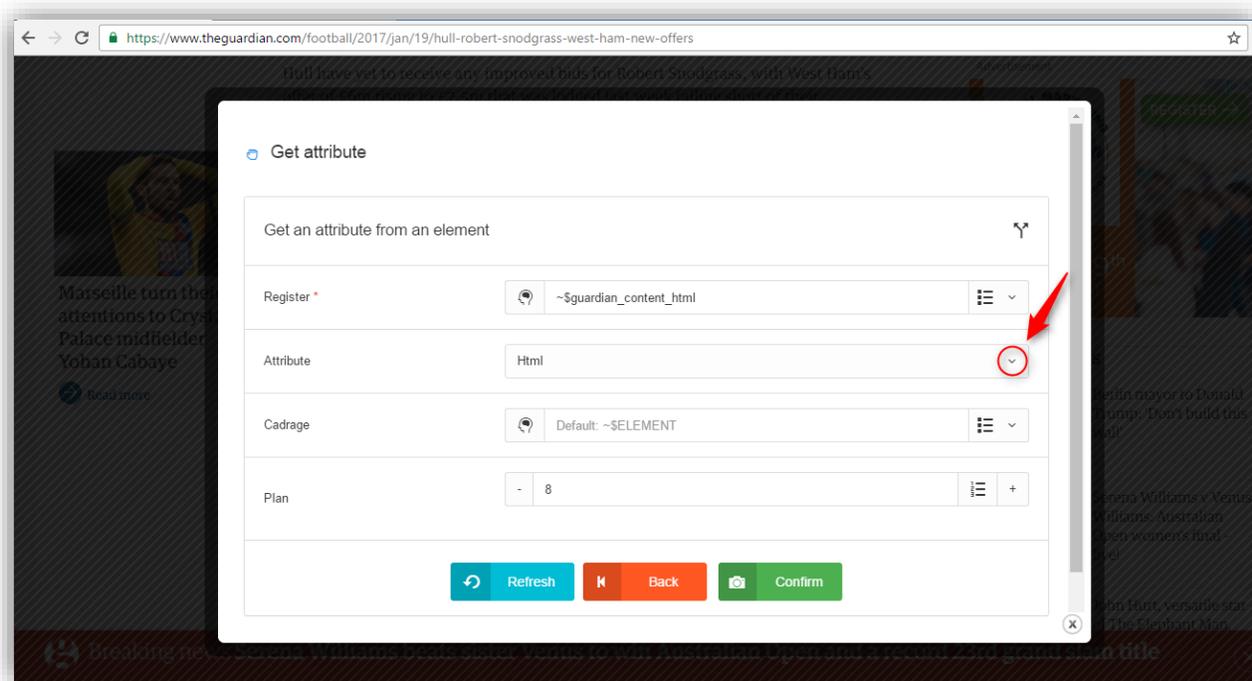
## کام نخست

بدون نیاز به خارج شدن از صفحه ی سامانه ، از کادر کشویی **Select an action** پلان **Get attribute** را انتخاب می کنیم.



## گام دوم

در فیلد **Register** نام رجیستر `~$guardian_content_html` را قرار می دهیم سپس از کادر کشویی **Attribute** گزینه ی **Html** را انتخاب می نماییم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .

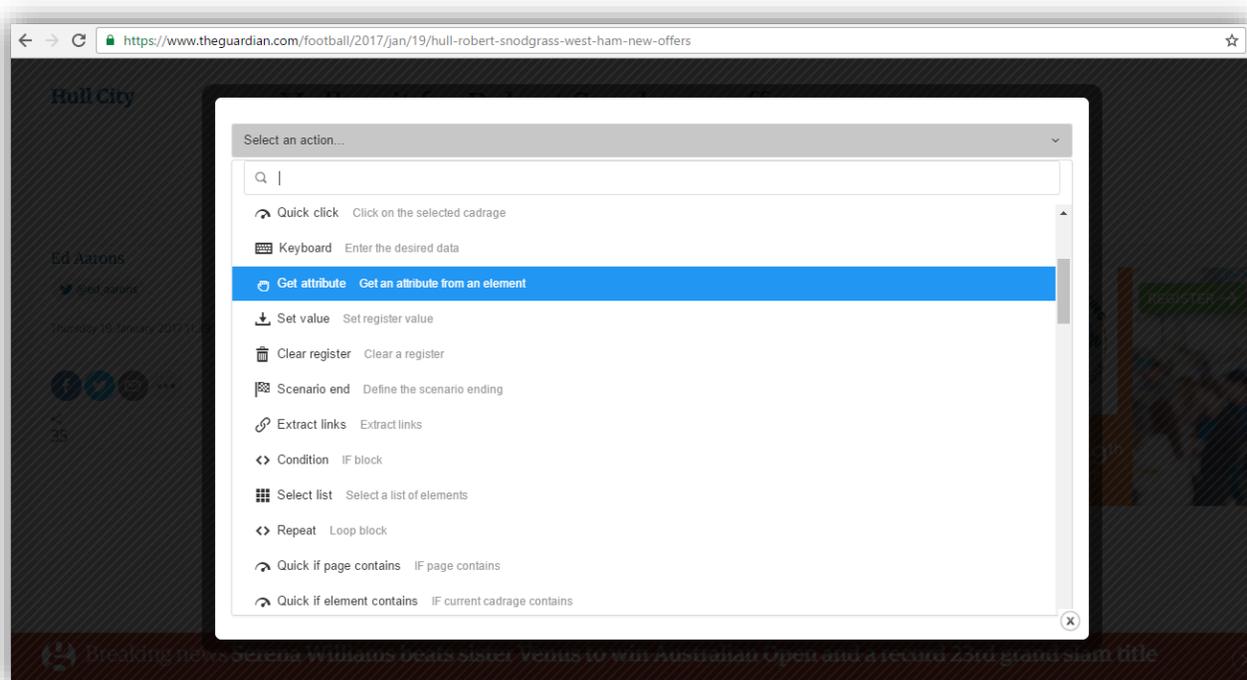


## پلان نه

در این قسمت می خواهیم ژلیوگراف به کمک پلان **Get attribute** متن خبر را به تنهایی درون یک رجیستر ذخیره کند .

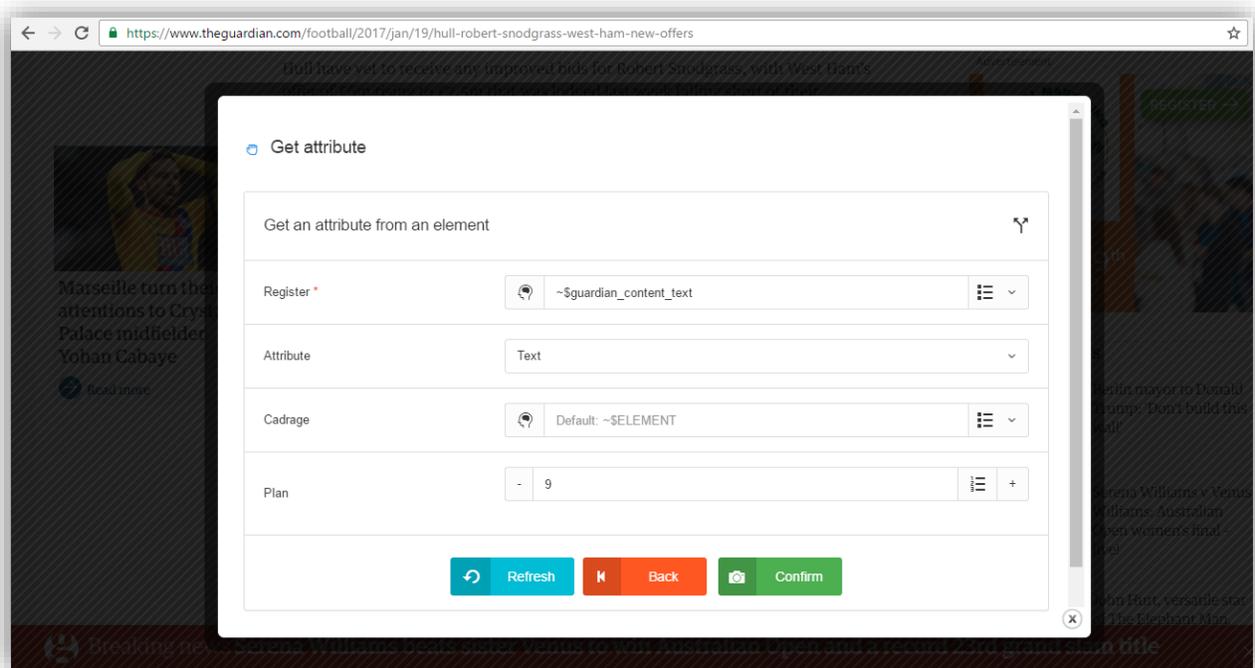
## کام نخست

بدون نیاز به خارج شدن از صفحه ی سامانه ، بار دیگر از کادر کشویی **Select an action** پلان **Get attribute** را انتخاب می کنیم .



## گام دوم

در فیلد **Register** نام رجیستر `~$guardian_content_text` را قرار می دهیم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .

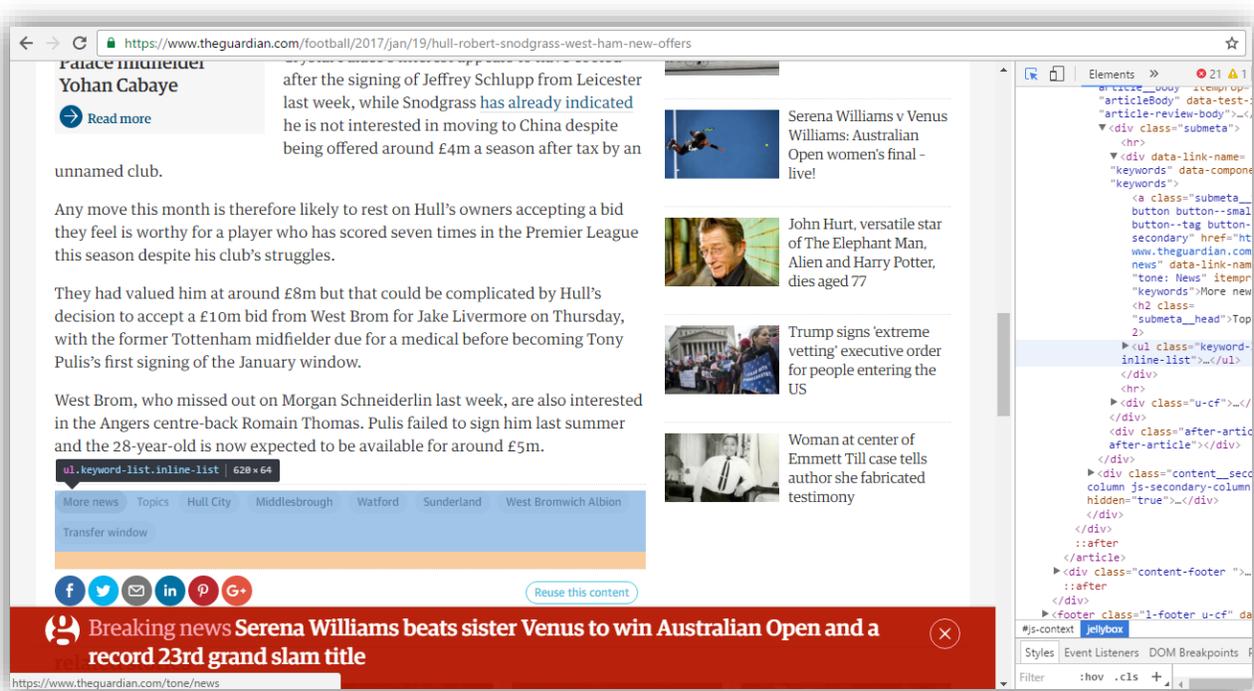


## پلان ده

در ادامه می خواهیم تا ژلیوگراف با استفاده از پلان **Select** کلید واژه های مربوط به خبر را انتخاب کند و در پلان بعد ، آن را درون یک رجیستر ذخیره نماید .

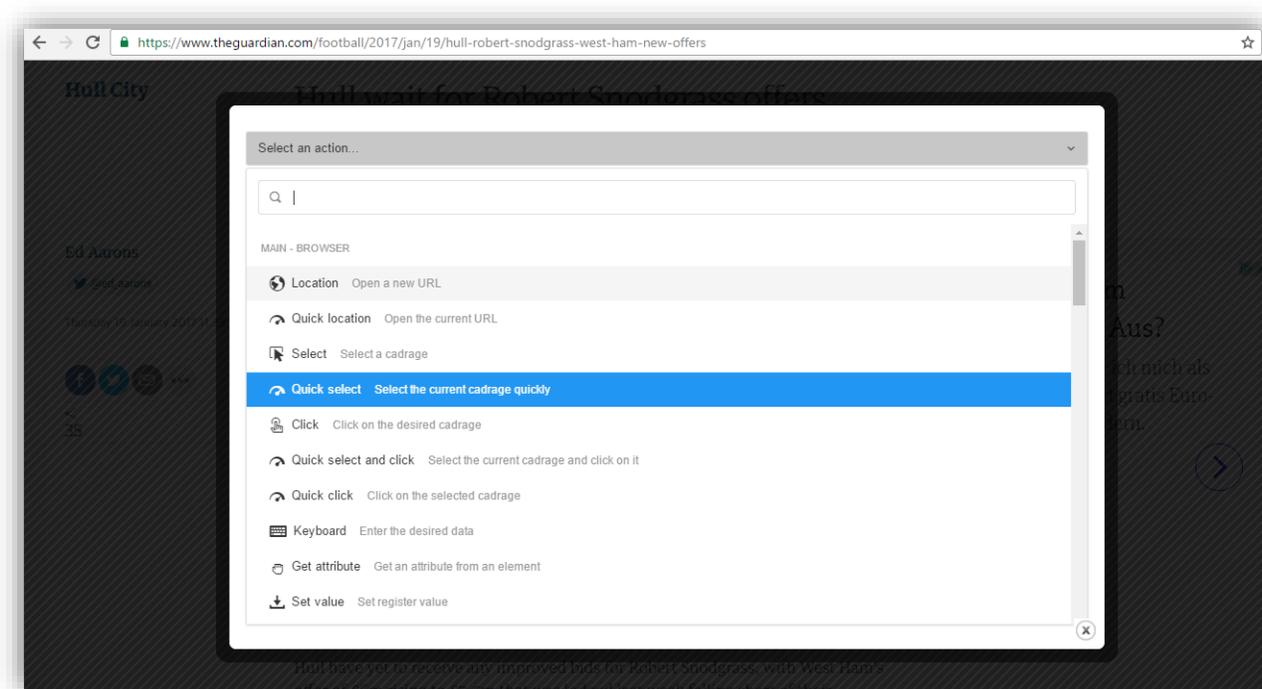
## گام نخست

پس از فشردن کلیدهای **Ctrl+Shift+C** اشاره گر ماوس را بر روی عبارات مربوط به کلیدواژه ی خبر قرار می دهیم و یک بار کلیک می کنیم تا صفحه ی سامانه گشوده شود .



## گام دوم

از کادر کشویی **Select an action** گزینه **Quick select** را انتخاب می کنیم .

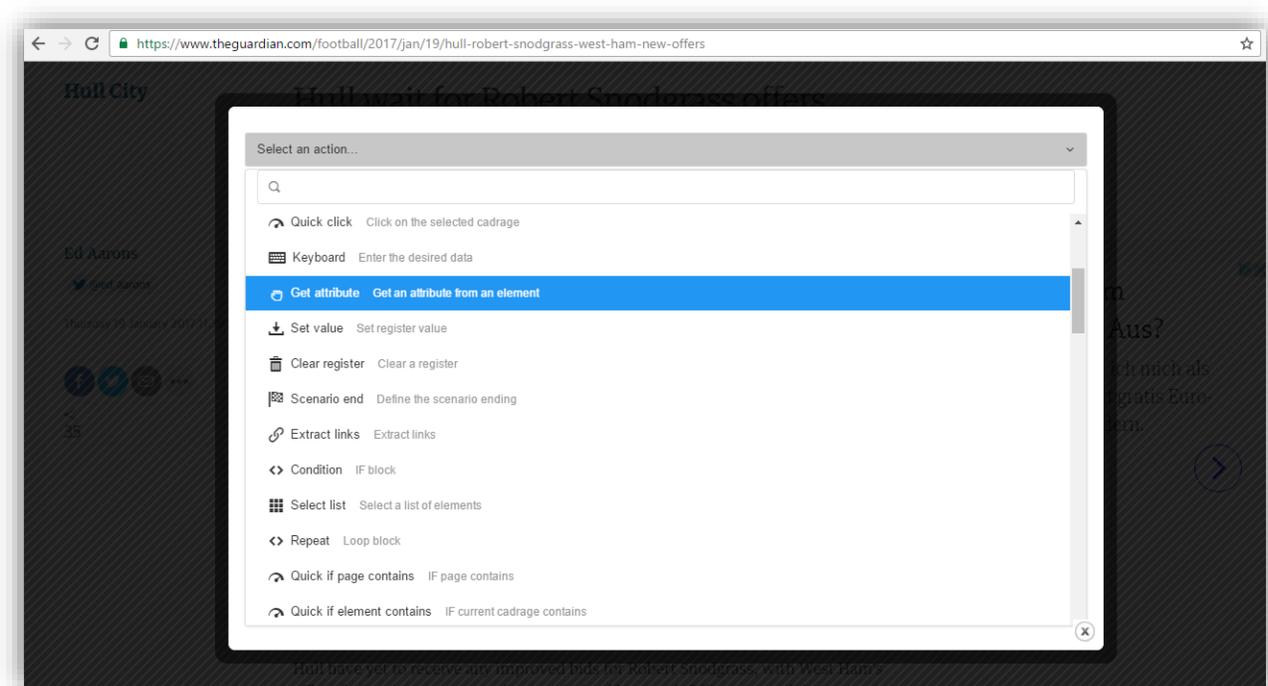


## پلان یازده

در این قسمت لازم است تا ژلیوگراف به کمک پلان **Get attribute** کلیدواژه های خبر را درون یک رجیستر ذخیره کند .

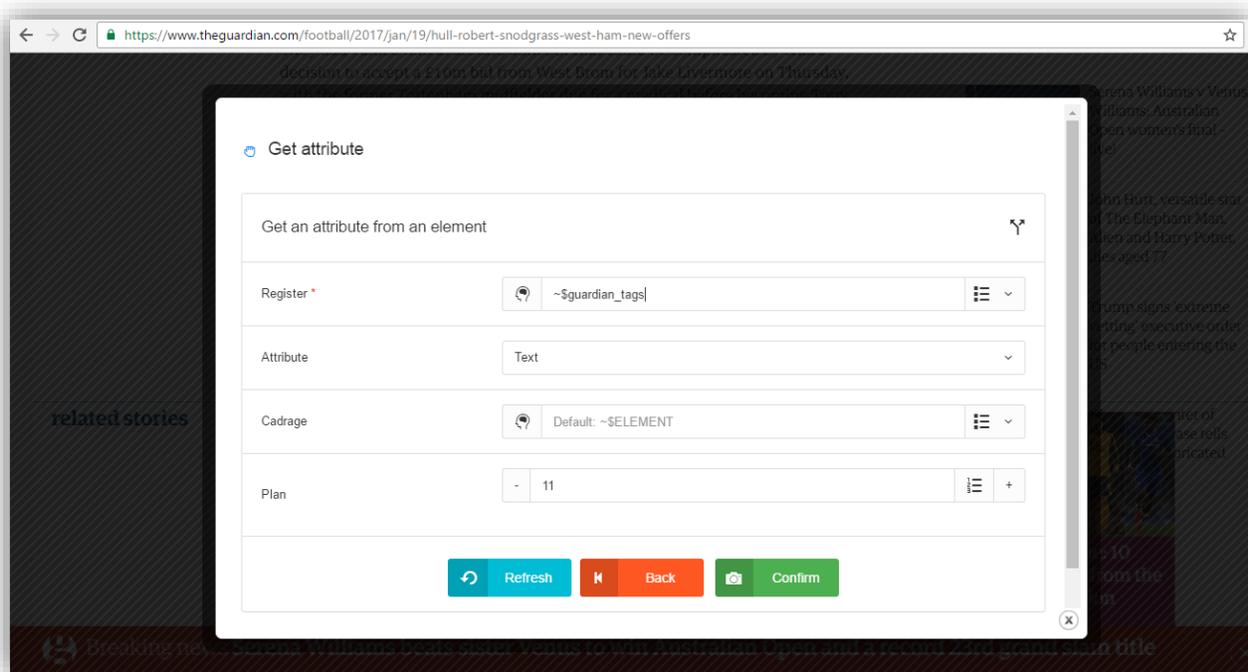
## کام نخست

بدون نیاز به خارج شدن از صفحه ی سامانه ، از کادر کشویی **Select an action** پلان **Get attribute** را انتخاب می کنیم.



## گام دوم

در فیلد **Register** نام رجیستر **~\$guardian\_tags** را قرار می دهیم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .



## پلان دوازده

در این قسمت لازم است چنانچه خبر دارای عکس شاخص باشد عکس شاخص دانلود شود و درون یک رجیستر قرار گیرد .  
در نظر داریم سناریو را با تعیین یک شرط اینطور ادامه دهیم که چنانچه در صفحه عکس شاخص وجود داشت ، با نام **index.png** دانلود شود .

## کام نخست

پس از فشردن کلیدهای **Ctrl+Shift+C** اشاره گر ماوس را بر روی عکس شاخص قرار می دهیم و یک بار کلیک می کنیم تا صفحه ی سامانه گشوده شود .

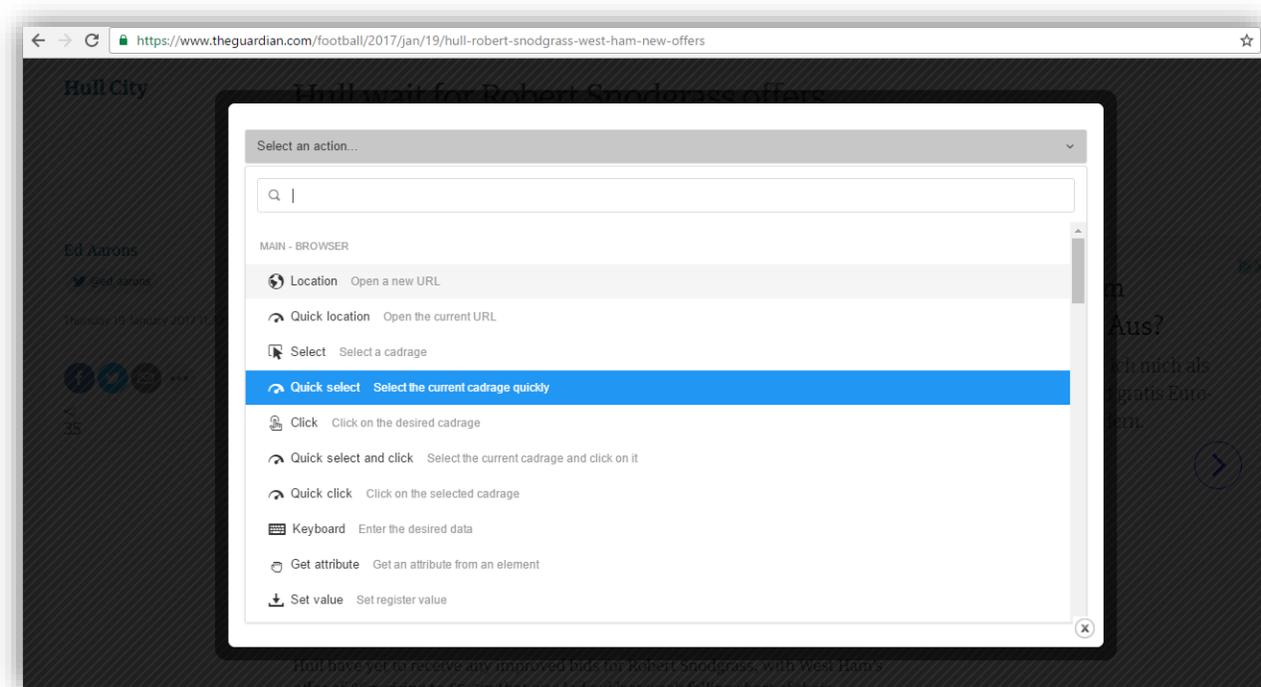
The screenshot shows a web browser displaying a news article from The Guardian. The article title is "Hull wait for Robert Snodgrass offers after rejecting West Ham's £7.5m bid". Below the title, there is a sub-headline: "Barnley, Watford, Middlesbrough and Sunderland all interested in Snodgrass as he looks to move to West Brom for £10m". The main image shows Robert Snodgrass in a Hull City kit. Below the image, there is a caption: "Robert Snodgrass refused to sign a new three-year contract at Hull last month. Photograph: Greig Cowie/BPI/REX/Shutterstock".

On the right side of the article, there is an advertisement for a "GLOBAL MARKETING Symposium PRESENTED BY LIOHBRIDGE". The ad includes a "REGISTER" button and the text "MARCH 9th FREE EVENT".

The browser's developer tools are open, showing the HTML source code for the advertisement image. The code includes an `<img>` tag with the following attributes: `class="max-responsive-img", itemprop="content", alt="Robert Snodgrass", src="https://i.guim.co.uk/image/media/1493cb37ae72a2f31b3be0559e8bd4e28084e558auto:8um=128fi=max94c2276c9eb61d6c1ccc07a465w" data-bbox="520 480 695 600".`

## گام دوم

از کادر کشویی **Select an action** گزینه ی **Quick select** را انتخاب می کنیم .

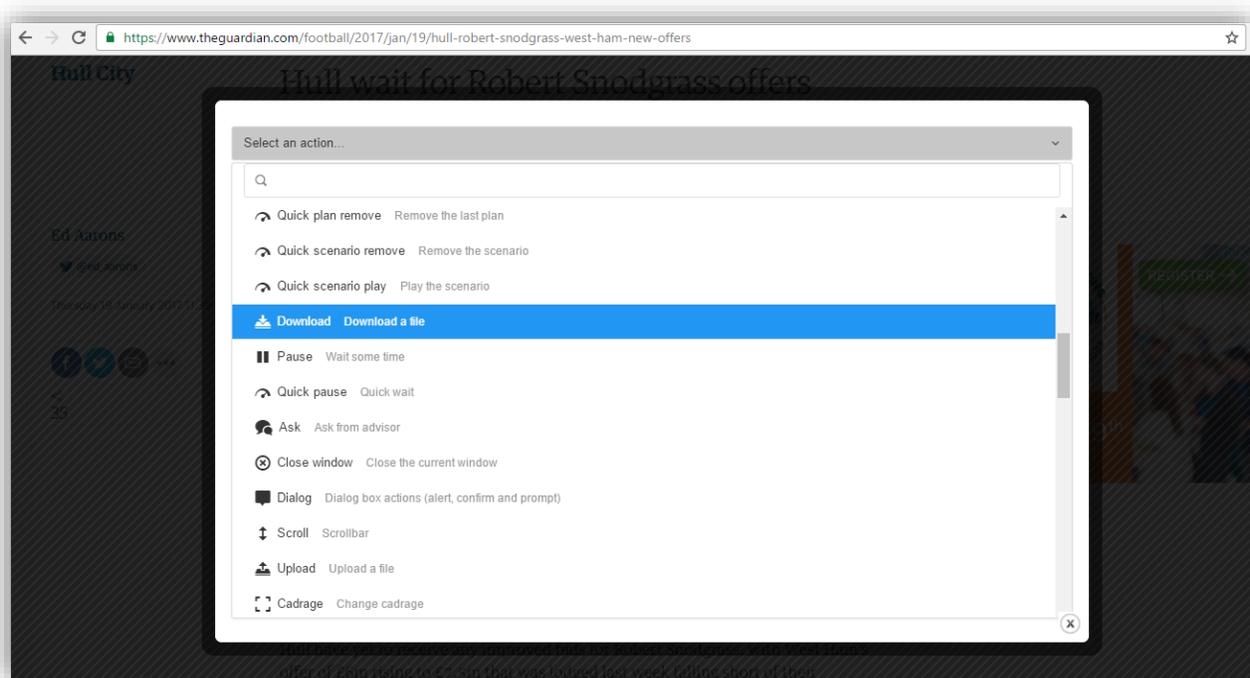


## پلان سیزده

در این مرحله ، ژلیوگراف با استفاده از پلان **Download** عکس انتخاب شده را دانلود می نماید .

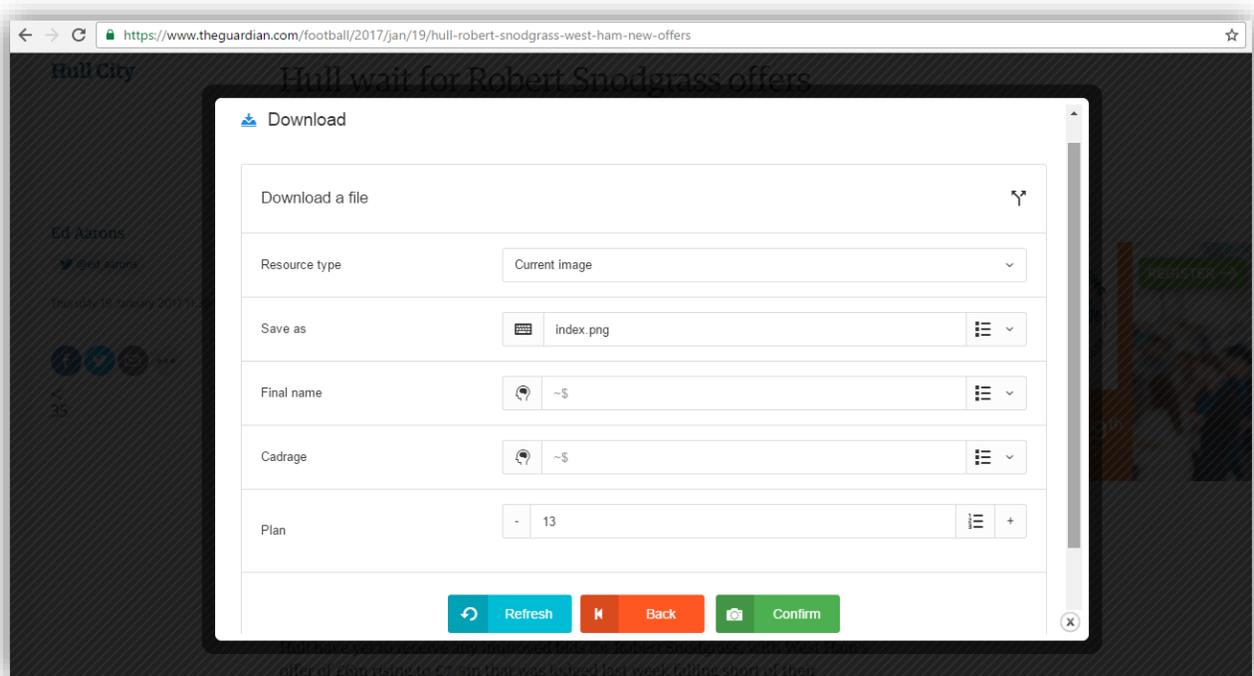
## گام نخست

بدون نیاز به خارج شدن از صفحه ی ژلیوگراف ، از کادر کشویی **Select an action** پلان **Download** را انتخاب می کنیم.



## گام دوم

از کادر کشویی **Resource type** گزینه **Current image** را انتخاب می کنیم و در فیلد **Save as** نامی که می خواهیم عکس شاخص با آن ذخیره شود ، وارد می نماییم و در انتها بر روی گزینه **Confirm** کلیک می کنیم .

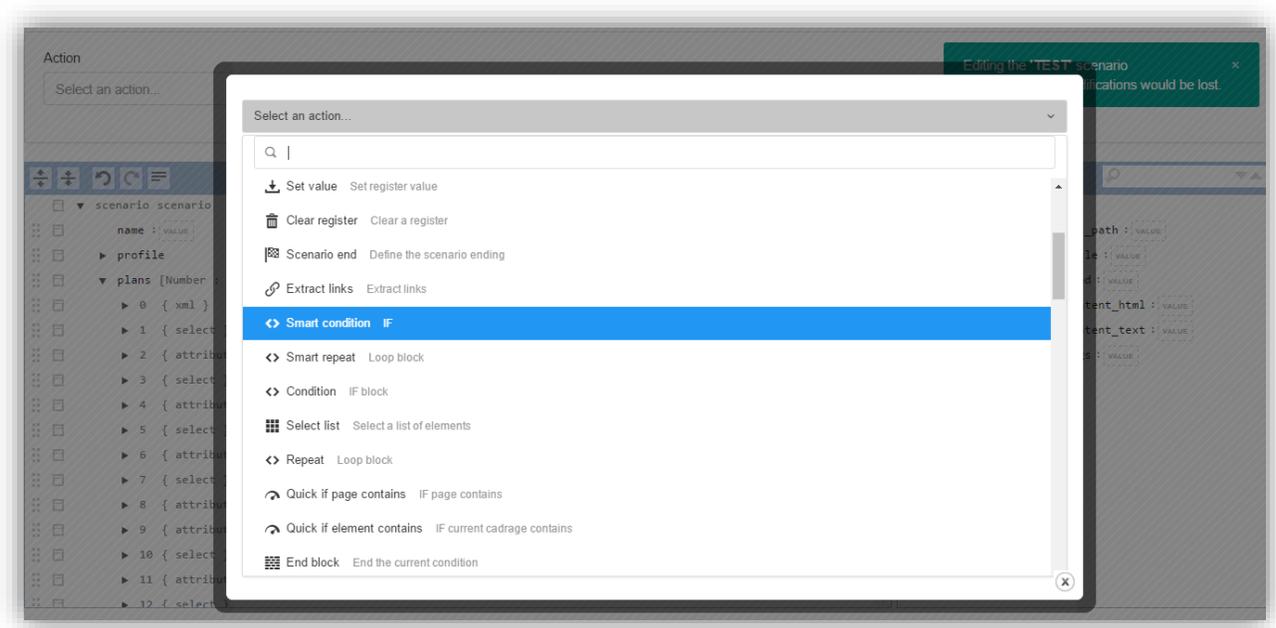


## پلان چهارده و پانزده

پس از بررسی متوجه شدیم که عکس شاخص دارای کلاس **maxed responsive-img** می باشد از آنجایی که ممکن است عکس شاخص در صفحه وجود نداشته باشد باید شرطی تعریف کنیم که اگر کلاس عکس انتخاب شده با عکس شاخص برابر باشد ، دانلود انجام گیرد به این منظور باید پلان های مربوط به دانلود عکس شاخص را بین پلان های **If** و **End** قرار دهیم ، پس برای تعیین شرط وارد صفحه ی **Scenario editor** می شویم .

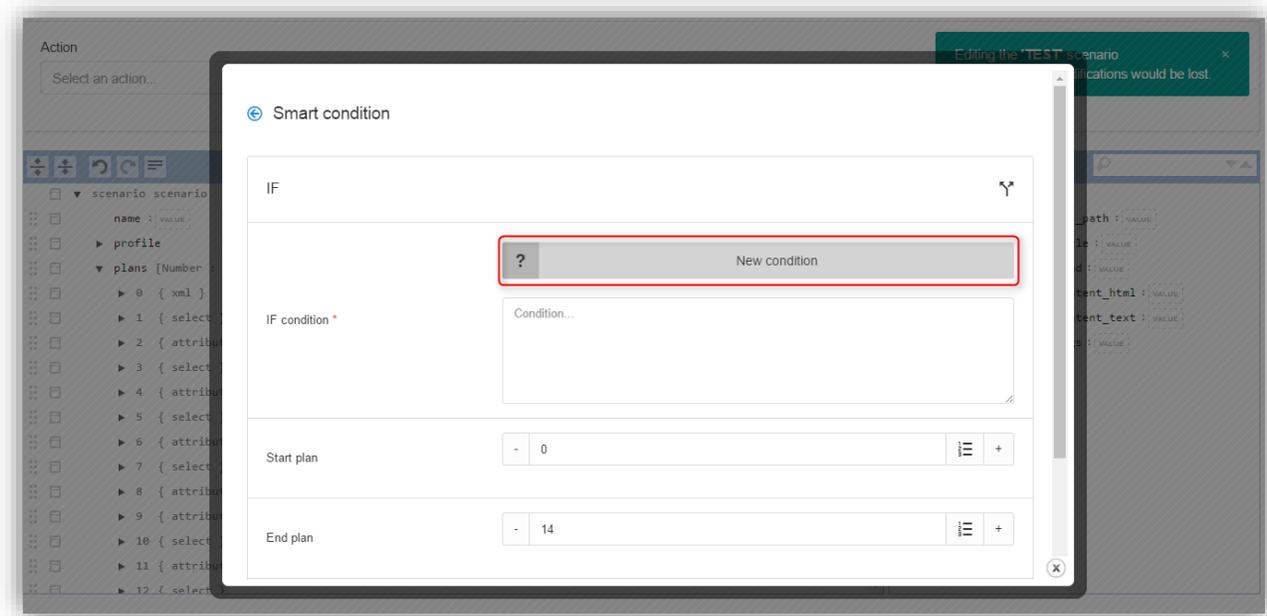
### کام نخست

پس از فشردن کلیدهای **Alt+J** از کادر کشویی **Select an action** گزینه ی **Smart condition** را انتخاب می کنیم.



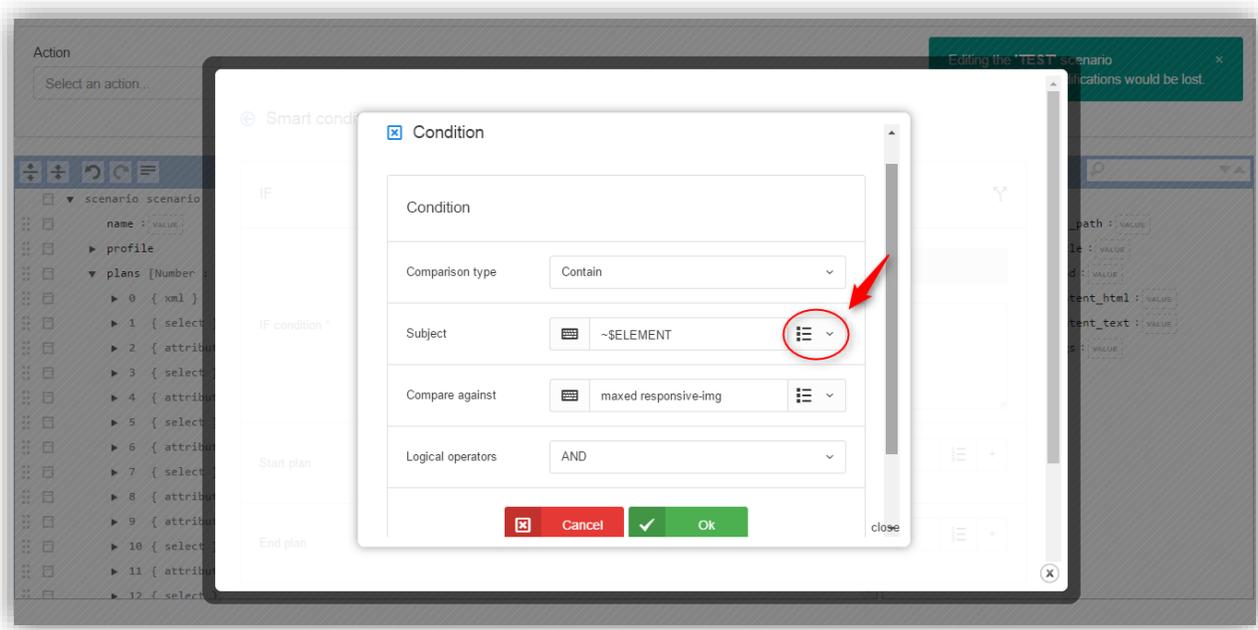
## گام دوم

بر روی کادر **New condition** یک بار کلیک می کنیم .



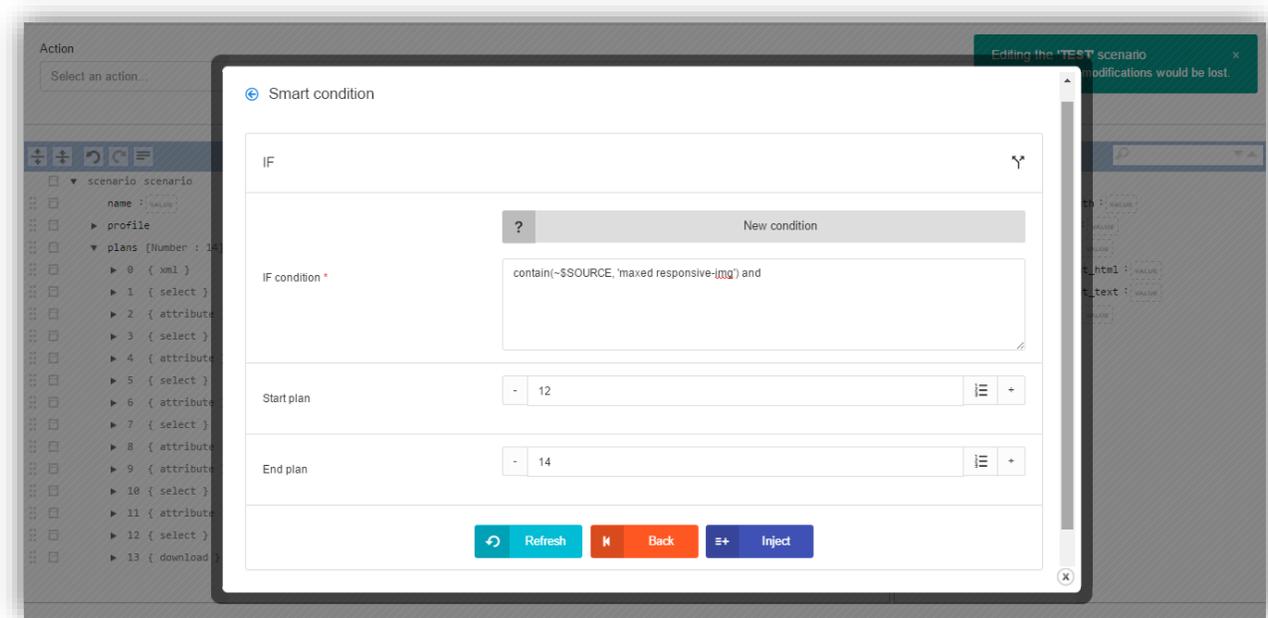
## گام سوم

از کادر کشویی **Subject** عبارت **~\$ELEMENT** را انتخاب می کنیم و در کادر **Compare against** نام کلاس عکس شاخص که برابر با **maxed responsive-img** می باشد را وارد می نماییم و در ادامه بر روی گزینه ی **OK** کلیک می کنیم.



## گام چهارم

پلان های دوازده و سیزده مربوط به دانلود عکس شاخص می باشد به منظور قرار دادن آن ها بین پلان های **If** و **End**، در فیلد **Start plan** عدد ۱۲ و در فیلد **End plan** عدد ۱۴ را قرار می دهیم و در انتها بر روی گزینه ی **Inject** کلیک می کنیم .

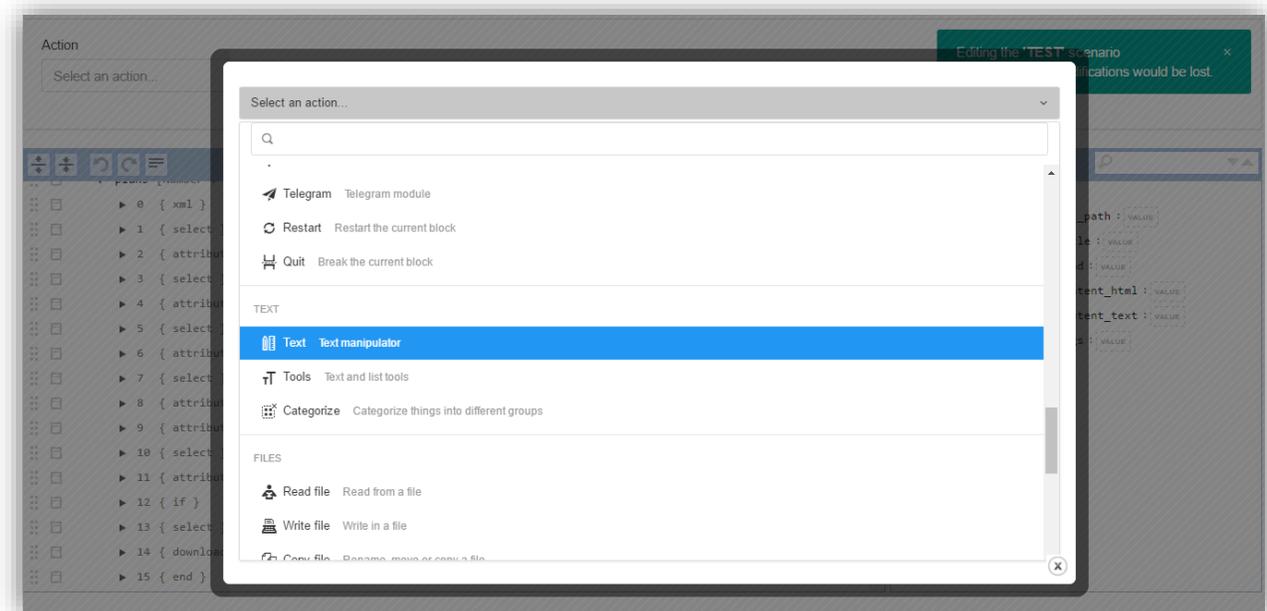


## پلان شانزده

ممکن است که در برخی از خبرها خلاصه وجود نداشته باشد در این گونه موارد می خواهیم به منظور خالی نبودن رجیستر مربوط به خلاصه ی خبر، پاراگراف اول متن خبر را درون رجیستر `$guardian_lead` قرار دهیم بدین منظور از پلان **Text manipulator** استفاده می کنیم .

## کام نخست

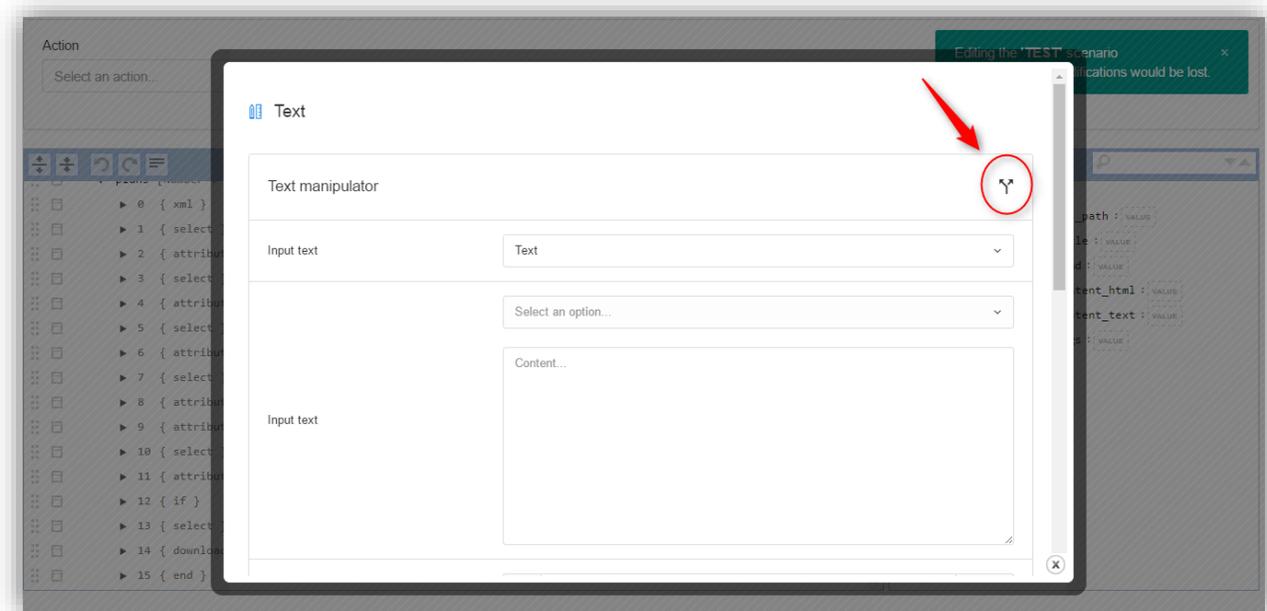
از کادر کشویی **Select an action** پلان **Text manipulator** را انتخاب می کنیم .



## گام دوم

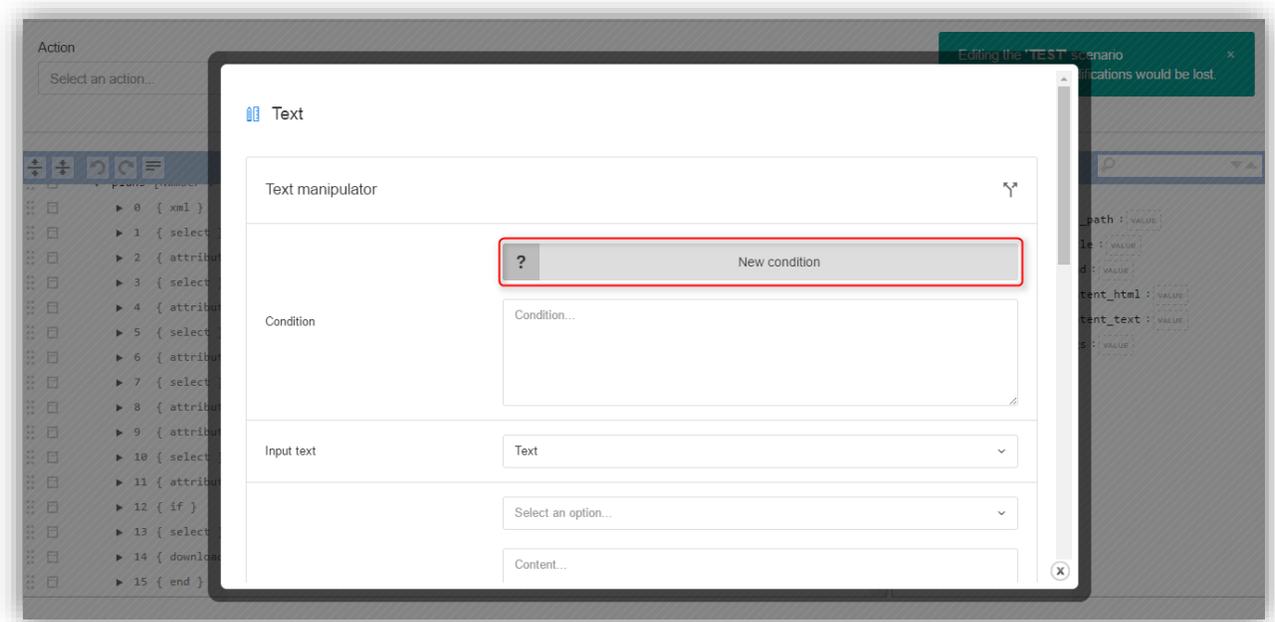
با توجه به اینکه می‌خواهیم این کار صرفاً زمانی انجام شود که خبر فاقد لید می‌باشد، باید قبل از هر اقدامی تعیین کنیم فقط زمانی که رجیستر `$guardian_lead` دارای مقدار نیست این پلان اجرا گردد.

بدین منظور بر روی علامت شرط یک بار کلیک می‌کنیم.



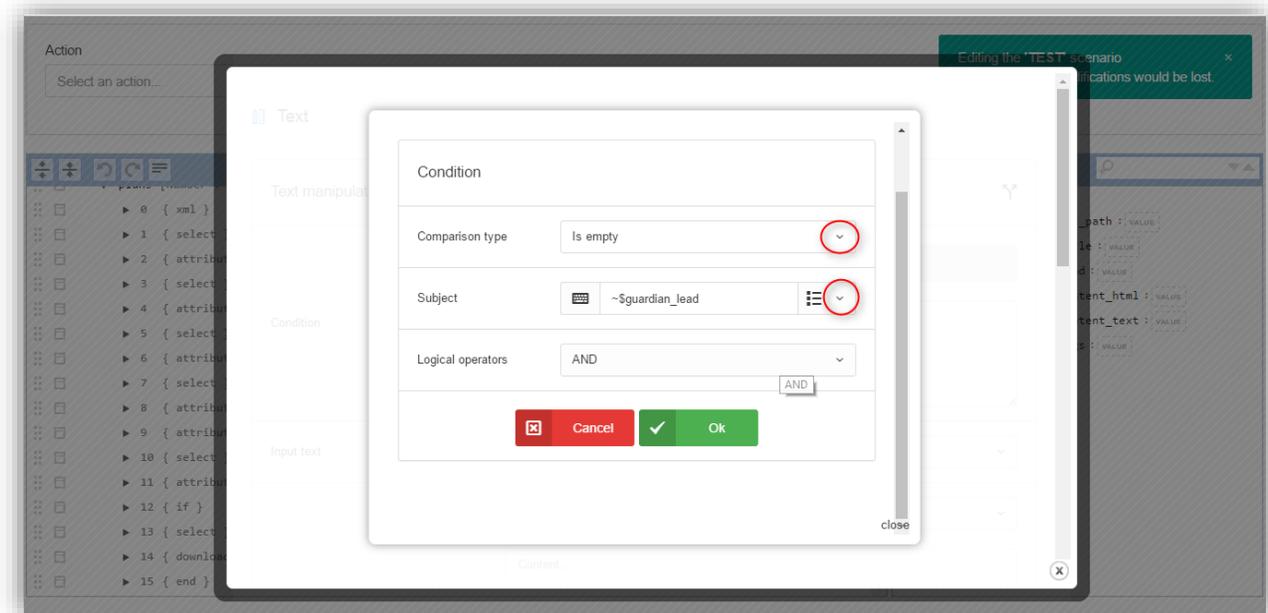
## گام سوم

کادر **New condition** را انتخاب می کنیم .



## گام چهارم

به کمک کادر کشویی **Comparison type** گزینه **Is empty** را انتخاب می نماییم و در ادامه از کادر کشویی موجود در فیلد **Subject** رجیستر **~\$guardian\_lead** را قرار می دهیم و بر روی گزینه **OK** کلیک می کنیم .



## گام پنجم

از کادر کشویی **Select an option** که در فیلد **Input text** قرار دارد ، رجیستر **~\$guardian\_content\_text** را انتخاب می کنیم .

رجیستر **~\$guardian\_lead** را به عنوان خروجی از کادر کشویی موجود در فیلد **Output register** انتخاب می نمایم تا اولین پاراگراف متن در این رجیستر قرار گیرد .

در ادامه به منظور انتخاب اولین پاراگراف متن لازم است ، از کادر کشویی **Action** گزینه ی **Get a paragraph** انتخاب گردد و در انتها بر روی گزینه **Confirm** کلیک شود .

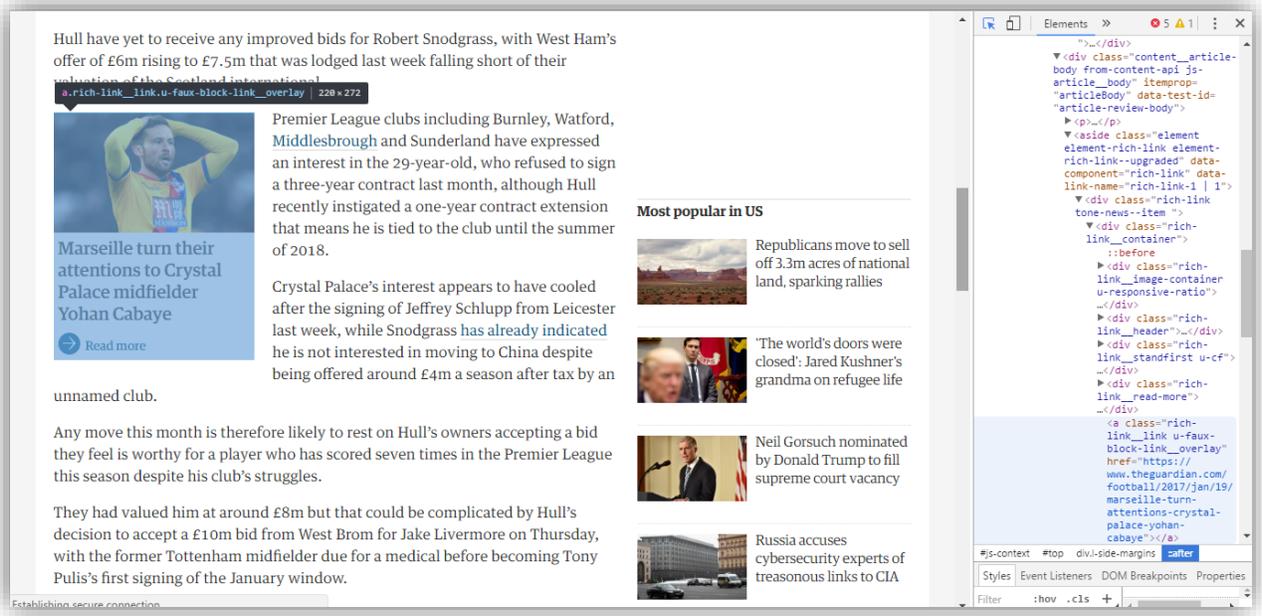
The screenshot shows the 'Text manipulator' interface with the following configuration:

- Condition:** A new condition is defined as `is_empty(~$guardian_lead) and`.
- Input text:** Set to `Text`.
- Input text (Action):** A dropdown menu is set to `Select an option...`, and the text field below contains `<~$guardian_content_text>`.
- Output register:** Set to `~$guardian_content_text`.
- Action:** Set to `Get a paragraph`.
- Paragraph num:** Set to `1`.
- Plan:** Set to `10`.

At the bottom, there are control buttons: Refresh, Back, Next step, Inject, and Confirm.

ممکن است در متن خبر کادر هایی وجود داشته باشد که مد نظر ما نباشد و بخواهیم آن ها را حذف کنیم .

ابتدا باید بدانیم برای حذف کادر چه خصوصیتی لازم می باشد . برای دانستن این موضوع کافی است در صفحه ی وبسایت **theguardian.com** ، پس از فشردن کلید های **Ctrl+Shift+C** اشاره گر ماوس را بر روی کادری که می خواهیم حذف شود ، قرار دهیم تا خصوصیات آن در کادر کوچک مشکی رنگ نمایش داده شود .



The screenshot shows a web browser displaying a news article on The Guardian website. The article is about Robert Snodgrass and Crystal Palace. The developer tool is open, showing the HTML structure of a rich-link element. The HTML code is as follows:

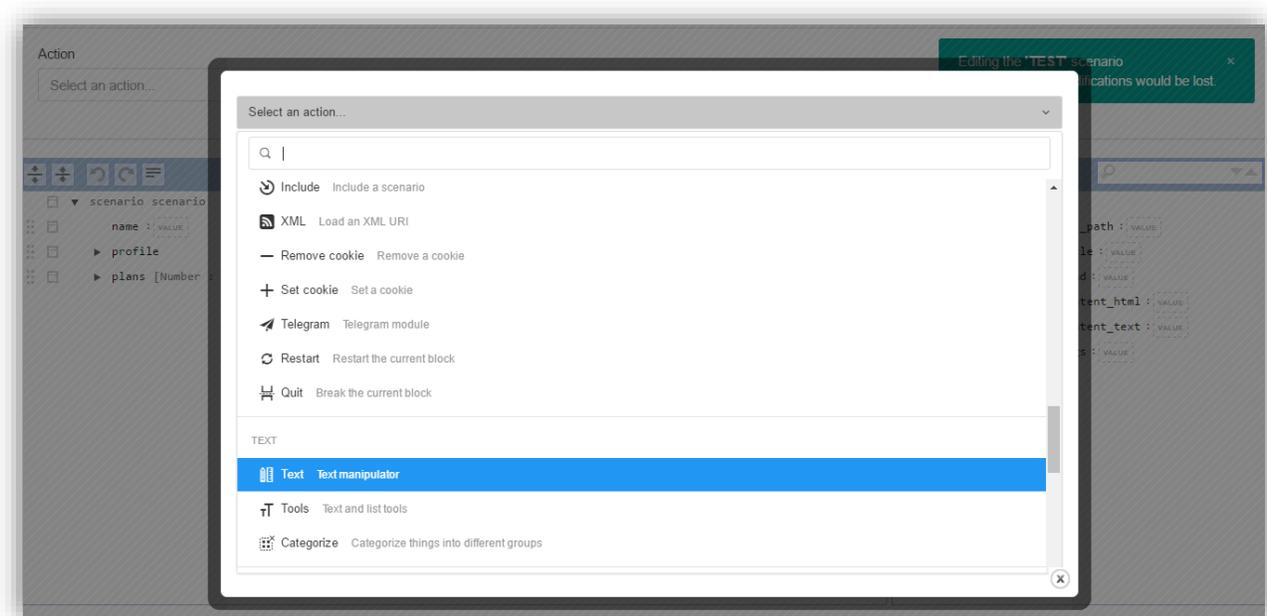
```
<div class="content_article-body from-content-api js-article-body" itemprop="articleBody" data-test-id="article-review-body">
  <p></p>
  <aside class="element element-rich-link element-rich-link-upgraded" data-component="rich-link" data-link-name="rich-link-1 | 1">
    <div class="rich-link tone-news-item">
      <div class="rich-link-container">
        <div class="rich-link_image-container u-responsive-ratio">
          </div>
        <div class="rich-link_header"></div>
        <div class="rich-link_standfirst u-ctf">
          <div class="rich-link_read-more">
            <a class="rich-link_link u-faux-block-link_overlay" href="https://www.theguardian.com/football/2017/jan/19/marseille-turn-attentions-crystal-palace-yohan-cabaye"></a>
          </div>
        </div>
      </div>
    </div>
  </div>
```

## پلان هفده

گاهی اوقات در میان متن خبر تعدادی کادر تبلیغاتی و یا خبری وجود دارد که باید حذف شوند و همچنین می خواهیم یک عبارت به ابتدای متن تمامی اخباری که استخراج می شوند اضافه کنیم . برای هر دو منظور یک بار از پلان **Text manipulator** استفاده می کنیم .

## گام نخست

ابتدا به صفحه **Scenario editor** باز می گردیم سپس از کادر کشویی **Select an action** پلان **Text manipulator** را انتخاب می کنیم .

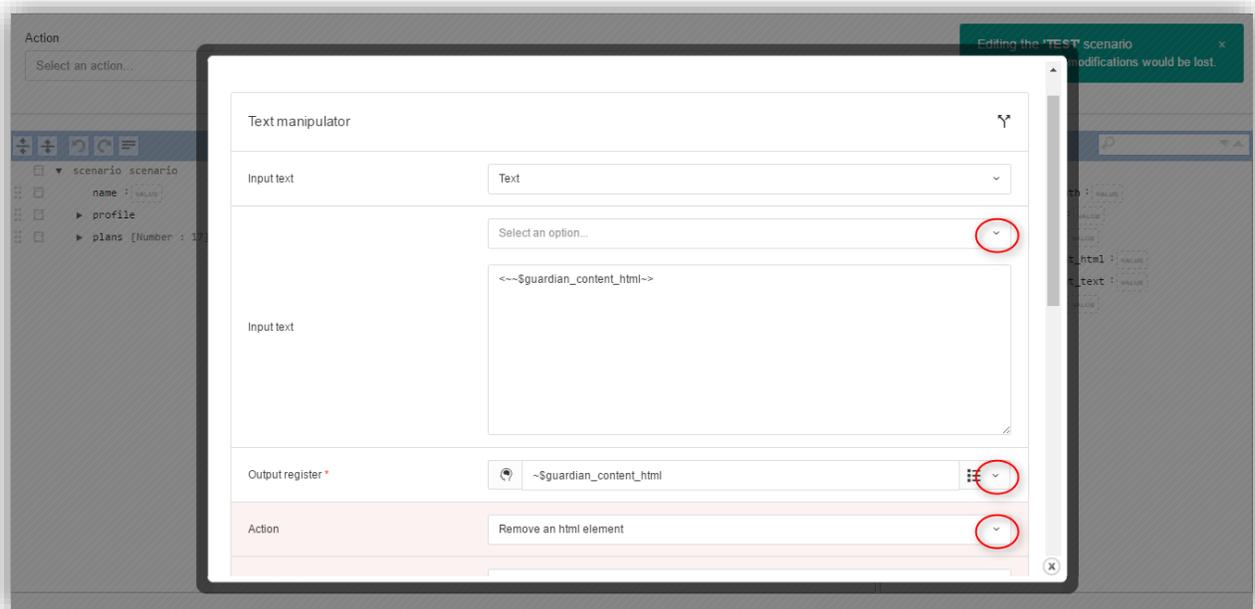


## گام دوم

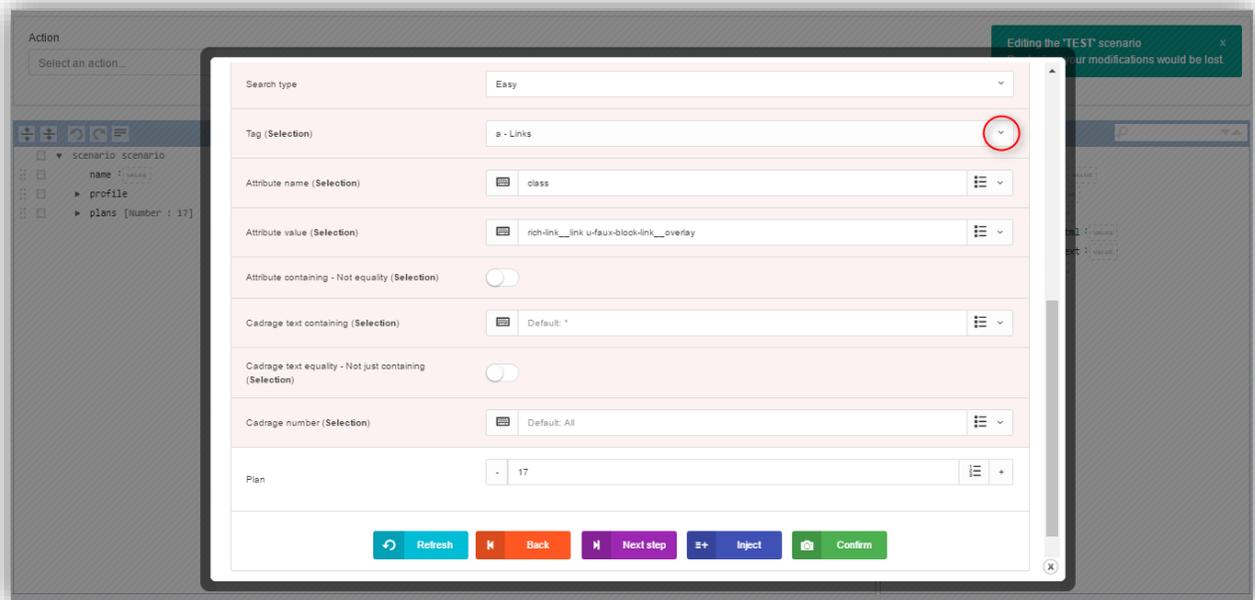
از کادر کشویی **Select an option** که در فیلد **Input text** قرار دارد ، رجیستر **~\$guardian\_content\_html** را انتخاب می کنیم .

از کادر کشویی موجود در فیلد **Output register** رجیستر **~\$guardian\_content\_html** را تعیین می کنیم که متن پس از تغییرات دوباره در همان رجیستر قرار گیرد .

از کادر کشویی **Action** گزینه ی **Remove an html element** را به منظور حذف کادر انتخاب می کنیم .

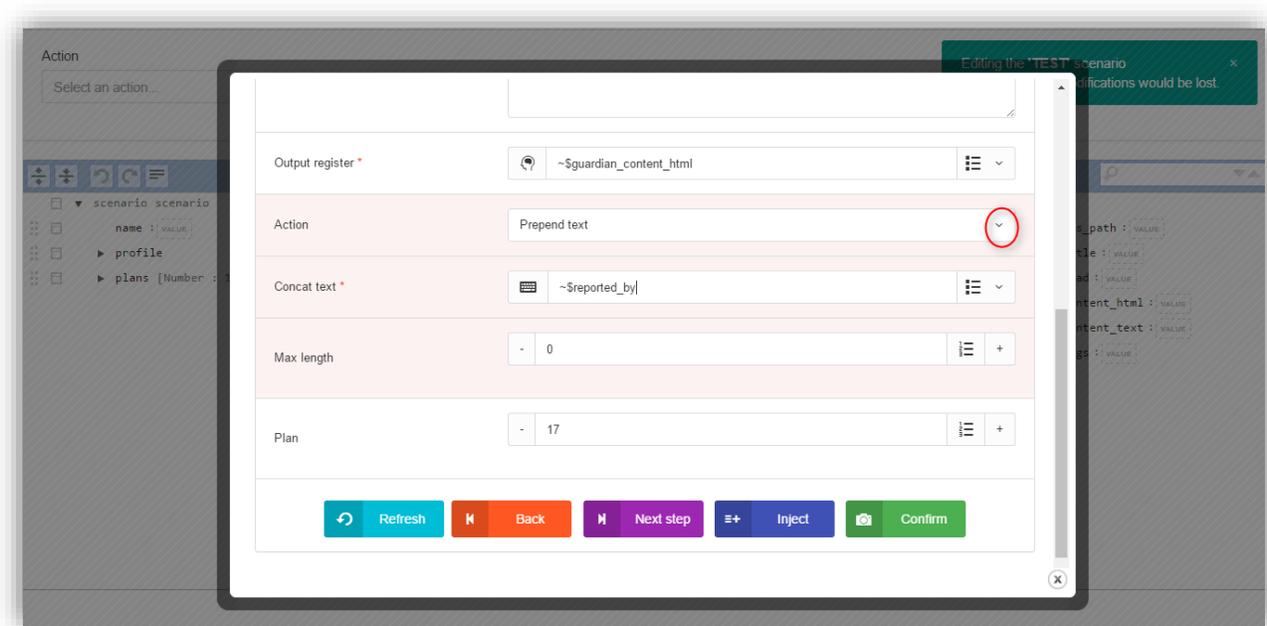


در ادامه لازم است از کادر کشویی موجود در فیلد **Tag** گزینه ی **a-links** را انتخاب نماییم و در فیلد **Attribute name** نیز عبارت **Class** را وارد کنیم . در انتها کلاس مورد نظر را که برابر با **rich-link\_\_link u-faux-block-link\_\_overlay** می باشد در فیلد **Attribute value** قرار می دهیم ، سپس جهت اضافه کردن عبارت مورد نظر به ابتدای متن بر روی گزینه ی **Next step** کلیک می کنیم .



## گام سوم

پس از انتخاب گزینه ی **Next step**، از کادر کشویی **Action** گزینه ی **Prepend text** را انتخاب می نمایم و در فیلد **Concat text** یک رجیستر قرار می دهیم تا در هر بار اجرای سناریو بتوانیم عبارتی که می خواهیم ابتدای متن قرار گیرد را تغییر دهیم و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم .

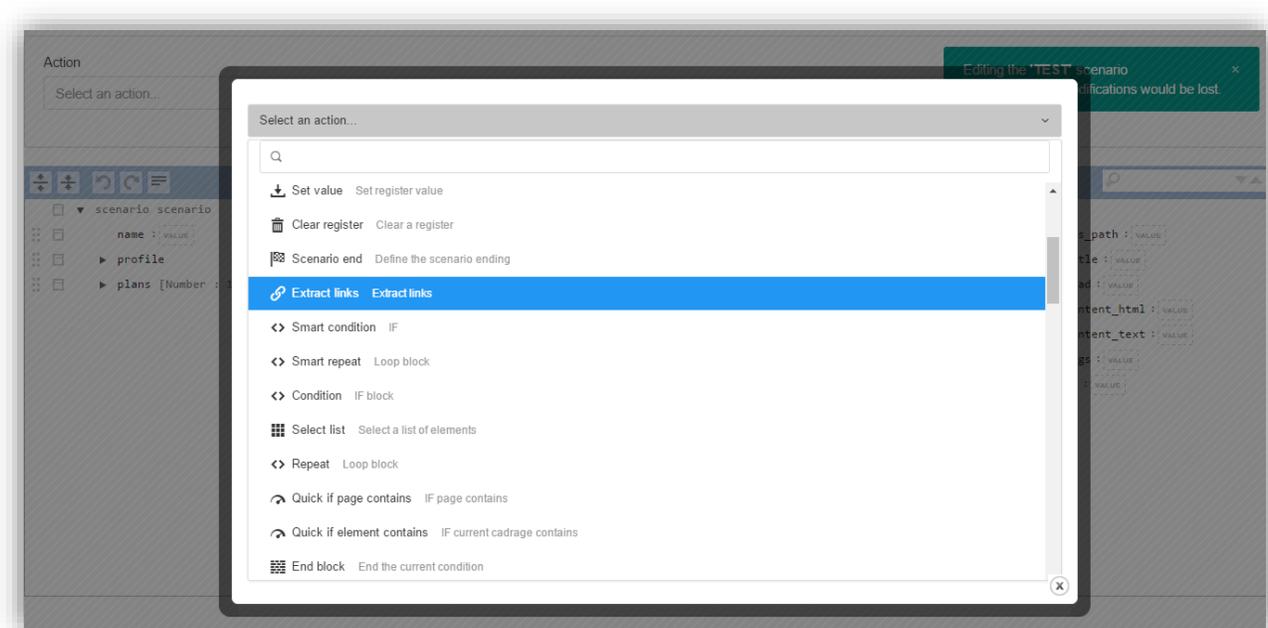


## پلان هجده

به منظور دانلود عکس های موجود در متن عکس توسط ژلیوگراف ، می خواهیم از پلان های **Extract links** و **Download** استفاده کنیم .

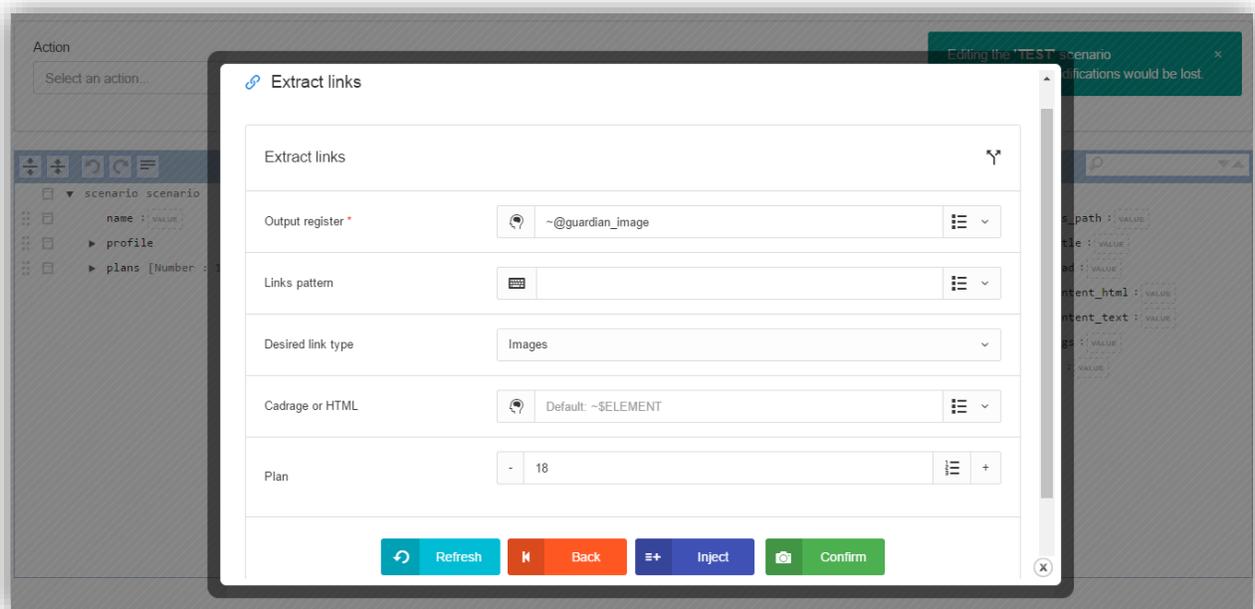
## گام نخست

از کادر کشویی **Select an action** پلان **Extract links** را انتخاب می کنیم .



## گام دوم

از آنجایی که ممکن است تعداد تصاویر موجود در متن بیش از یکی باشد پس لازم است قبل از نام متغیر علامت `@` قرار گیرد پس در فیلد **Output register** عبارت `@guardian_image` را نوشته و در انتها بر روی گزینه ی **Confirm** کلیک می کنیم. بعد از اجرای این پلان آدرس تمامی تصاویر استخراج شده درون این رجیستر ذخیره می گردد.

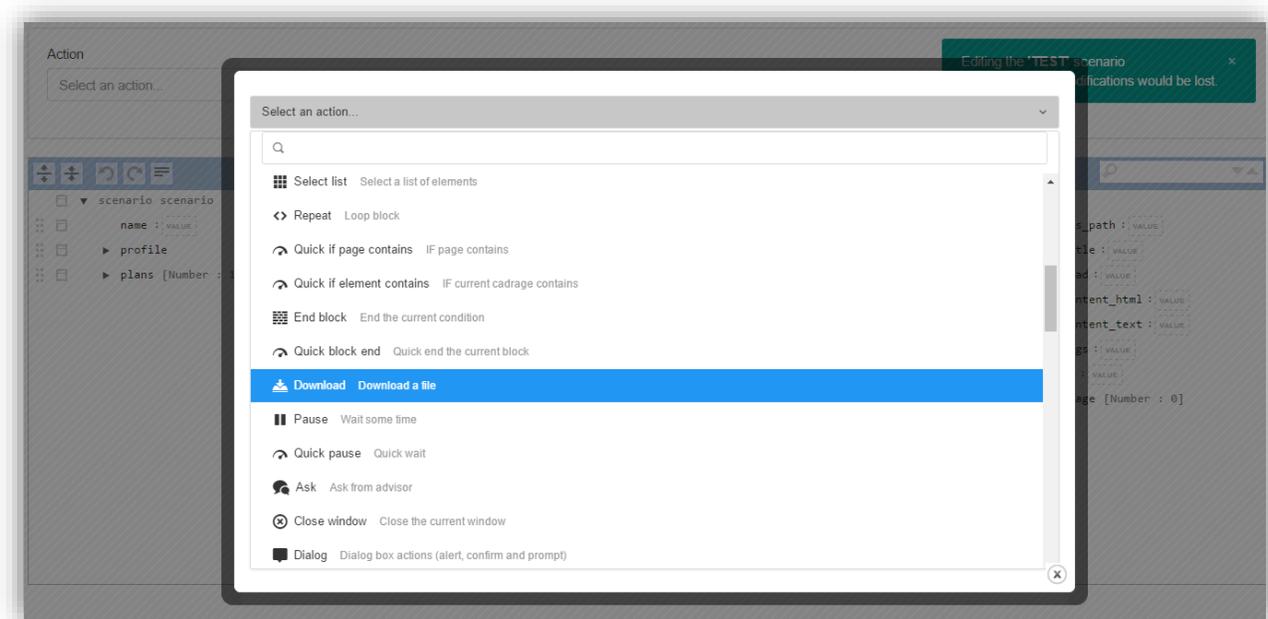


## پلان نوزده

در پلان قبل آدرس تمامی تصاویر موجود در متن درون یک رجیستر قرار گرفتند. اینک لازم است تا عکس ها توسط ژلیوگراف دانلود شوند.

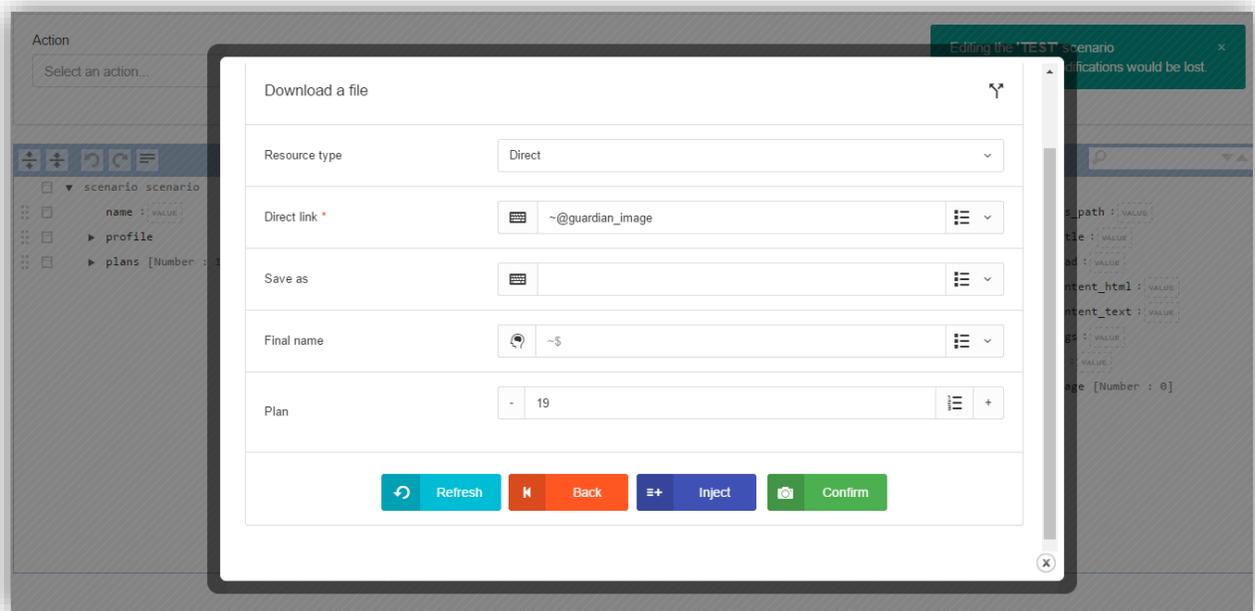
## گام نخست

از کادر کشویی **Select an action** پلان **Download** را انتخاب می کنیم.



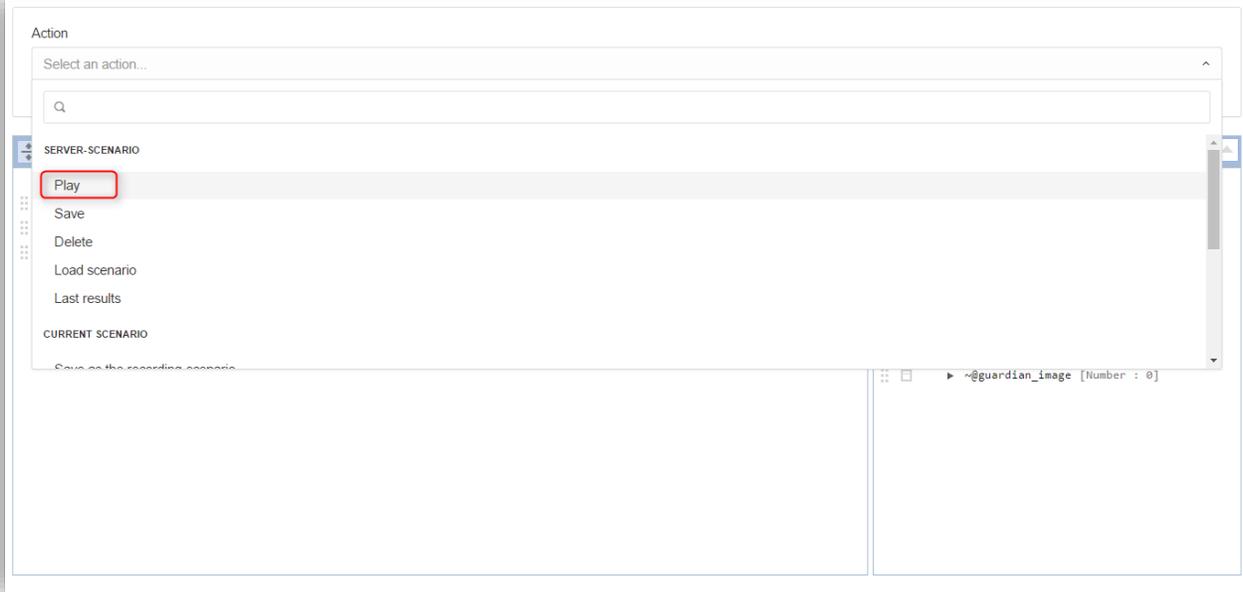
## گام دوم

در این مرحله کافی است رجیستر `@guardian_image` را در فیلد **Direct link** قرار دهیم و در انتها بر روی گزینه ی **Confirm** کلیک نماییم. لازم به ذکر است در اینجا هر تصویر با نام واقعی خود در سایت مبدأ ذخیره می گردد .

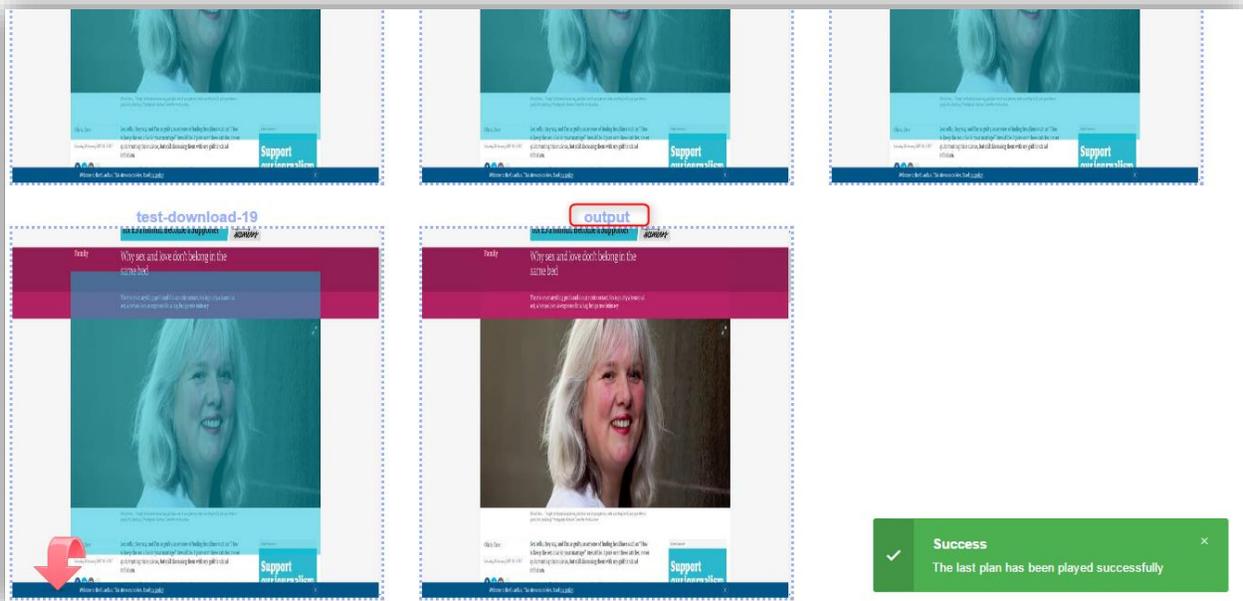


## اجرای سناریو

جهت اجرای سناریو پس از خارج شدن از سامانه ، از کادر کشویی **Select an action** گزینه ی **Play** را انتخاب می کنیم .



پس از پایان اجرای سناریو ، یک بار روی کادر **output** کلیک می کنیم تا نتایج حاصل از اجرای سناریو را مشاهده نماییم .

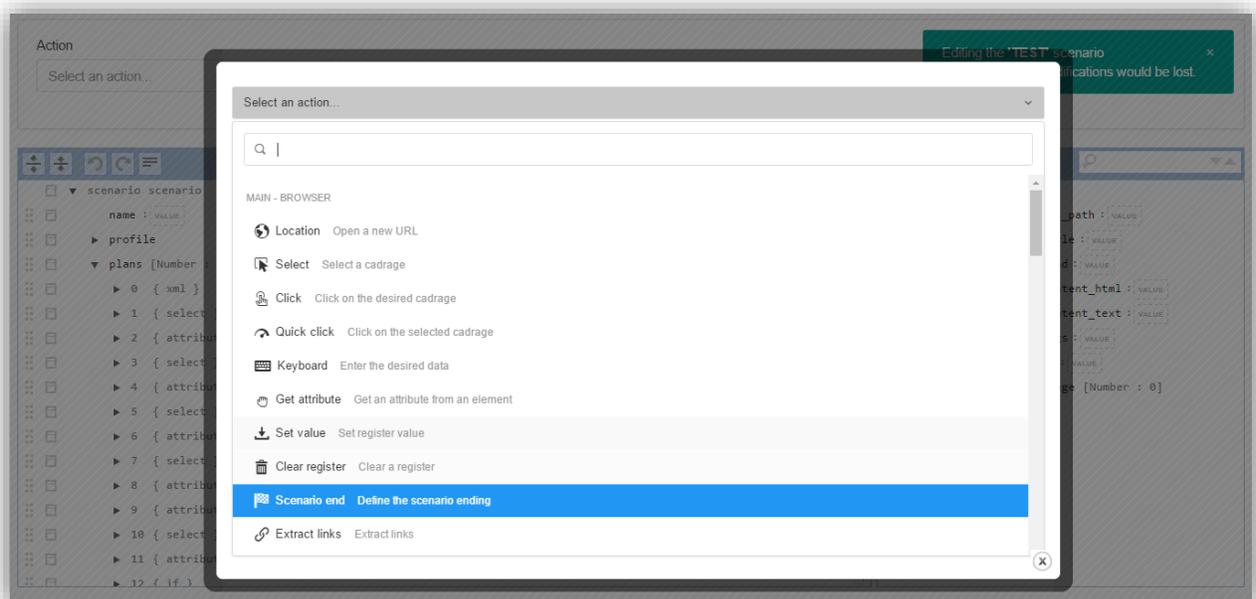


همانطور که در خروجی نیز مشاهده می شود ، قسمت های مختلف خبر در رجیستر های مربوط به خود قرار گرفته اند و نیز به ابتدای متن خبر عبارت **Jelliograph** اضافه شده است .

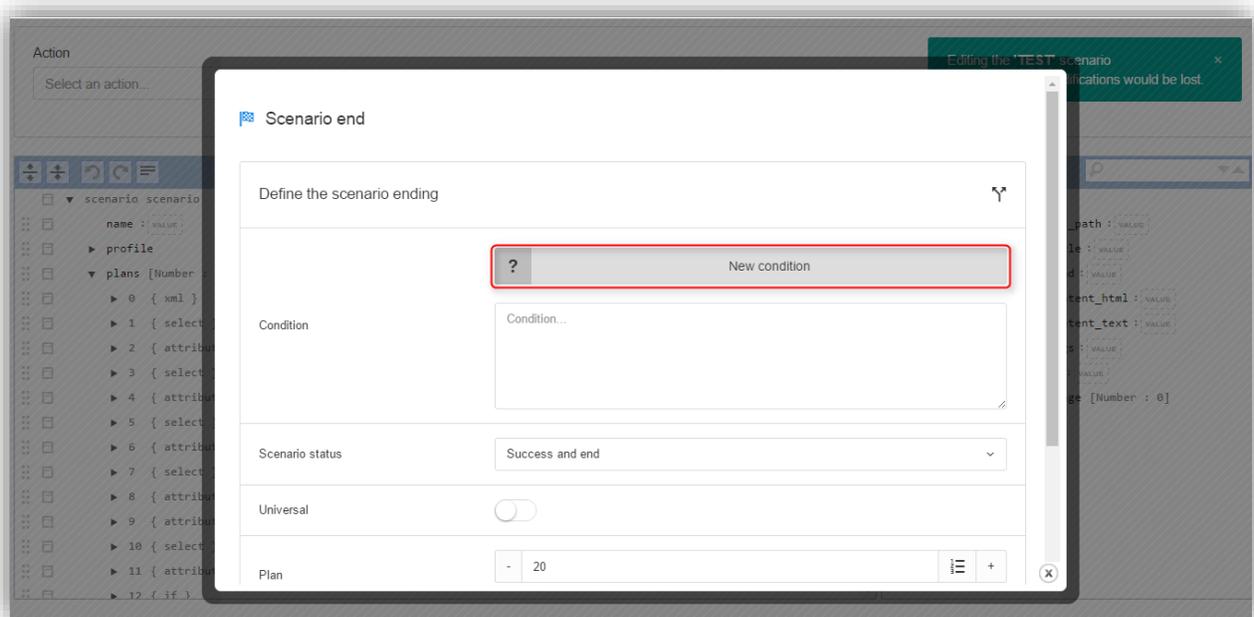
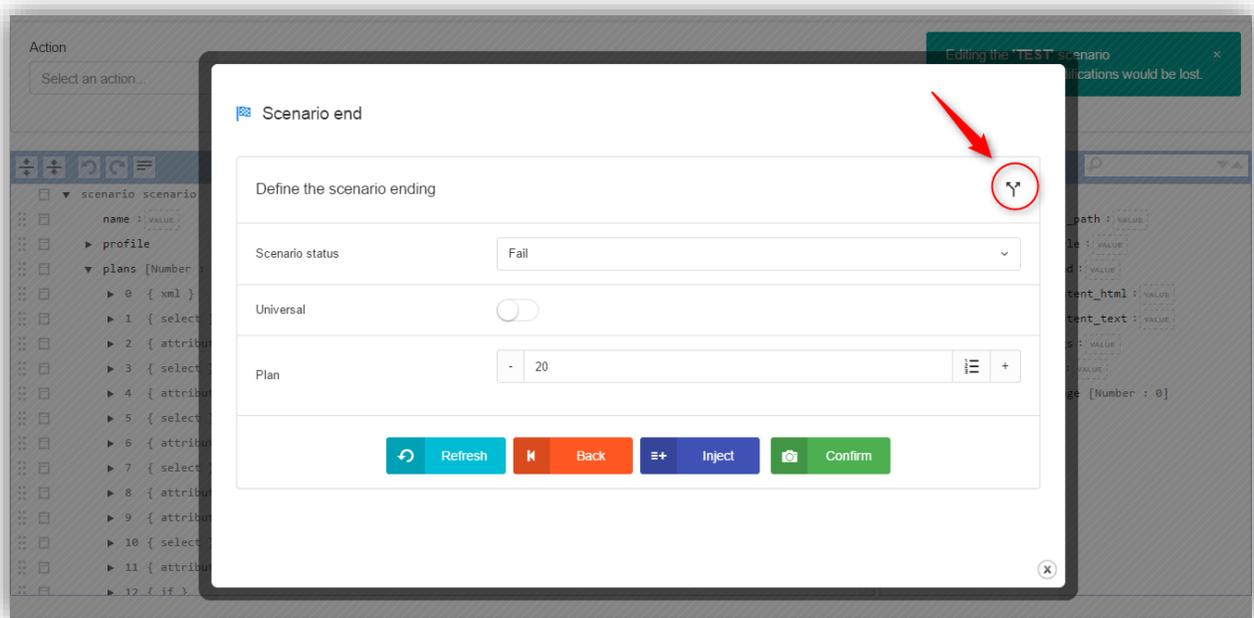
The image displays three screenshots of a web browser showing the HTML source code of an article from the Guardian. The first screenshot shows the main body text of the article, which discusses the complexities of sex, love, and relationships. The second screenshot shows the metadata and header information, with red circles highlighting 'Family Sex Relationships' and 'Jelliograph'. The third screenshot shows the footer and navigation elements, with red arrows pointing to 'Guardian News Path', 'Guardian Title', 'Campaign', and 'Guardian Lead'.

## بهبود سناریو

در وبسایت **theguardian.com** صفحاتی وجود دارد که تنها شامل یک ویدیو می شود و این صفحات مد نظر ما نمی باشد. طبق بررسی ها متوجه شدیم که در آدرس **URL** این صفحات کلمه ی **Video** وجود دارد پس می توان با قرار دادن پلان **Scenario end** بعد از پلان **XML** تعیین کرد که در صورت انتخاب چنین صفحاتی از اجرای پلان های بعدی جلوگیری شود. بدین منظور در صفحه ی **Scenario editor** پس از فشردن کلید های **Alt+J** از کادر کشویی **Select an action** پلان **Scenario end** را انتخاب می کنیم .

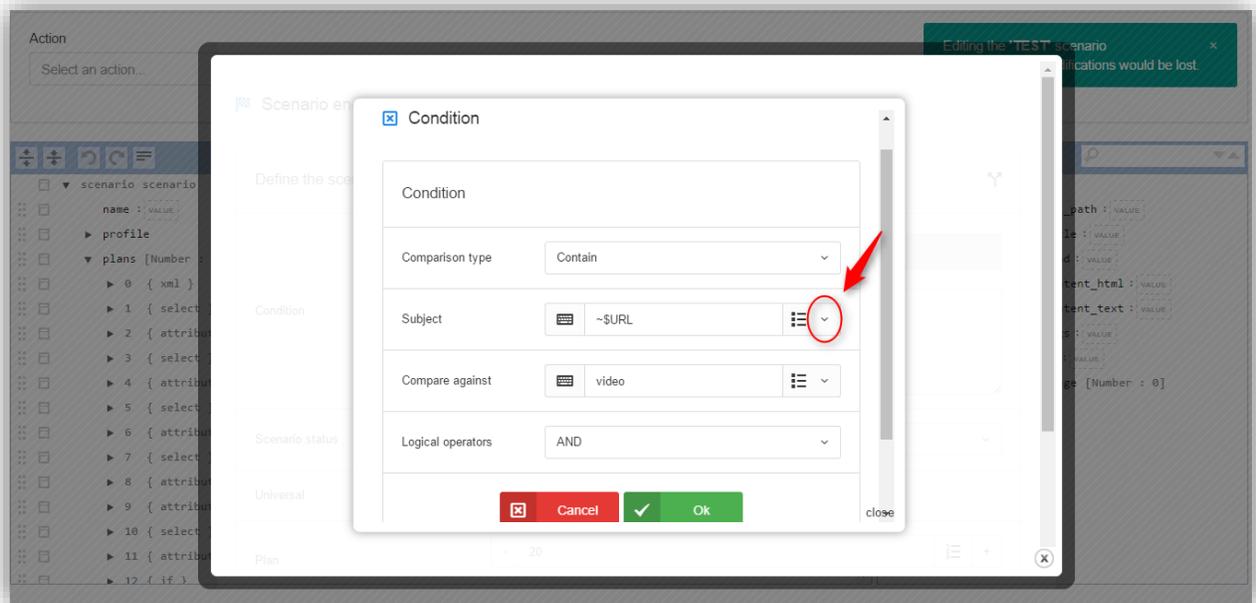


در ادامه بر روی علامت تعیین شرط کلیک می کنیم .



کادر **New condition** را انتخاب می کنیم .

به کمک کادر کشویی **Subject** گزینه ی **~\$URL** را انتخاب می کنیم و در فیلد **Compare against** عبارت **video** را قرار می دهیم و در پایان بر روی گزینه ی **OK** کلیک می کنیم .



در ادامه از کادر کشویی **Scenario status** گزینه ی **Success and end** را انتخاب می کنیم و در فیلد **Plan** عدد 1 را قرار می دهیم و در انتها بر روی گزینه ی **Inject** کلیک می کنیم .

