Power Automate for desktop CSSセレクタで Web操作が快適に



みなさん、こんにちは

名前: Fukuyori 或は Spumoni (カクテルの名前が由来)

出現場所:

Power BI勉強会、 おうじゃさんといっしょ RPA COMMUNITY Japan Microsoft Learn User Groupなど Qiita、noteも書いてます









今回の内容

Qiitaを参照してください。

https://bit.ly/3sVnLyc

今回の内容

01 cssセレク ターとは 02

Webページからデータを抽出する

03

セレクター、 結合子、フィ ルター 04

Yahoo!路線 情報のデータ 取得

CSSでデザイン定義を行う際、

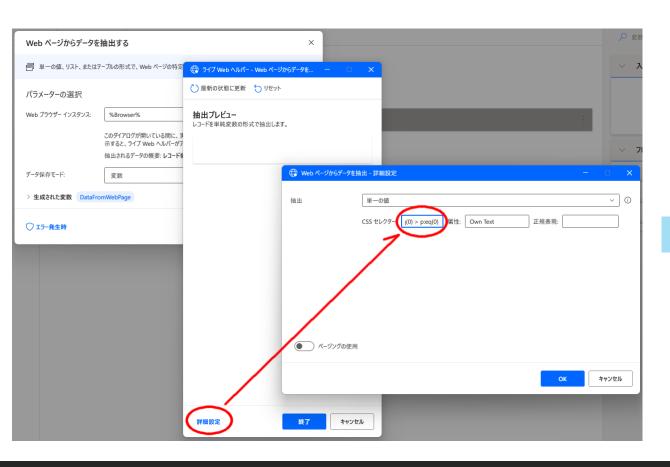
定義する場所を指定する記述方法です。



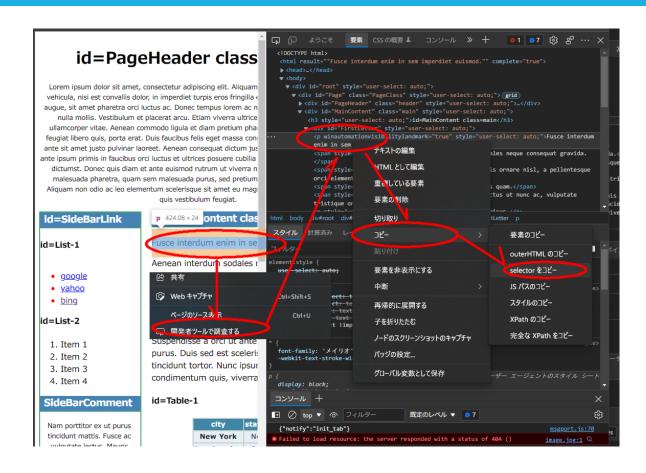
https://developer.mozilla.org/ja/docs/Web/CSS/CSS_Selectors



http://www.jquerystudy.info/reference/selectors/



html > body > div:eq(0) > div > div:eq(1) > div:eq(0) > p:eq(0)



#FirstLetter > p.bold-01

Power Automate for desktopは、最上部(html)から全てのセレクタを記述するが、Microsoft Edgeでは場所が特定できる部分を抜き出して記述している。

```
<!DOCTYPE html>
 3 (1) <html>
      <head>
        <title>HTML sample</title>
       <style type="text/css"> ···
121
        </style>
      </head>

<u>Shody></u>

124
      g<div id="root">
125
         <div id="Page" class="PageClass">
126 >
          <div id="PageHeader" class="header"> ···
139
          </div>
140
       <div id="MainContent" class="main">
141
142
           <h3>id=MainContent class=main</h3>
143
         <div id="FirstLetter">
144
            145
            span>Aenean interdum sodales neque consequat gravida.</span>
```

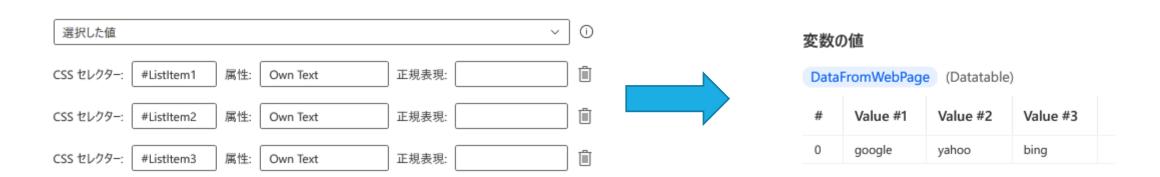
```
① ② ③ ④ ⑤ ⑦
html > body > div:eq(0) > div > div:eq(1) > div:eq(0) > p:eq(0)
```

```
<!DOCTYPE html>
      <html>
       <head>
        <title>HTML sample</title>
        <style type="text/css"> ···
 6 >
121
         </style>
       </head>
       <body>
124
         <div id="root">
125
          <div id="Page" class="PageClass">
           <div id="PageHeader" class="header"> ···
126 >
139
           </div>
140
141
           <div id="MainContent" class="main">
142
            <h3>id=MainContent class=main</h3>
143
          (1)<div id="FirstLetter">
            class="bold-01"> Fusce interdum enim in sem imperdiet euismod.
44
145
              <span>Aenean interdum sodales neque consequat gravida.</span>
```

```
① ②
#FirstLetter > p.bold-01
```



名称	説明
未定義	データを取得しません
単一の値	1つの値を取得します
選択した値	セレクタを複数指定して取得
リスト	リスト形式で一取得
テーブル	テーブル形式で一取得
HTMLテーブル全体	tableタグの内容をテーブル形式で取得



追加の CSS セレクターの指定

 抽出
 リスト
 > ①

 基本 CSS セレクター:
 #Table-1 > tbody > tr
 ①

 CSS セレクター:
 th
 属性:
 Own Text
 正規表現:

変数の値

DataF	romWebPage	(Datatable)
#	結果	
0	New York	
1	Los Angeles	
2	Chicago	
3	Miami	
4	Dallas	
5	Philadelphia	

テーブルや、リストで抽出する とき、基本CSSセレクターの指定す る場所は、繰り返しが発生する場 所です。

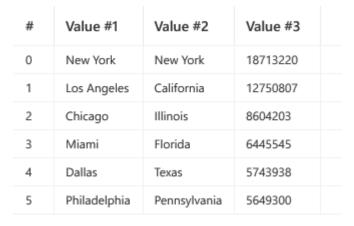
そのあと、CSSセレクターで各項 目の内容を指定してやります。



テーブル 0 \checkmark 1 CSS セレクター: #Table-1 > tbody > tr 属性: 正規表現: Ŵ CSS セレクター: Own Text Ŵ 正規表現: CSS セレクター: 属性: Own Text td:eq(0) Ŵ 属性: Own Text 正規表現: CSS セレクター: td:eq(1) 追加の CSS セレクターの指定

変数の値

DataFromWebPage (Datatable)







変数の値

DataFromWebPage (Datatable)

#	city	statename	population	density
0	New York	New York	18713220	10715
1	Los Angeles	California	12750807	3276
2	Chicago	Illinois	8604203	4574
3	Miami	Florida	6445545	5019
4	Dallas	Texas	5743938	1526
5	Philadelphia	Pennsylvania	5649300	4554

セレクター

名称	説明
全称セレクター	すべての要素
要素型セレクター	要素名を指定
クラスセレクター	class属性を指定
IDセレクター	id属性を指定
属性セレクター	属性を指定

例

名称	値	書き方
要素	a	a
クラス	urlclass	.urlclass a.urlclass
ID	link	#link a#link
属性	href id class	<pre>[href] a[id="link"] a[href^="https"][href*=".com"]</pre>

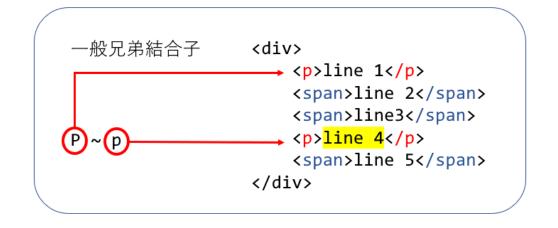
優先度は、

IDセレクター > 要素型セレクター > クラスセレクター = 属性セレクター

※IDセレクターは、原則として1ページ上でユニークになっています。原則に従っているページなら、IDが設定されている場所へ間違いなく飛ぶことができます。ただし、あくまで原則なので、IDが重複しているページもあります。

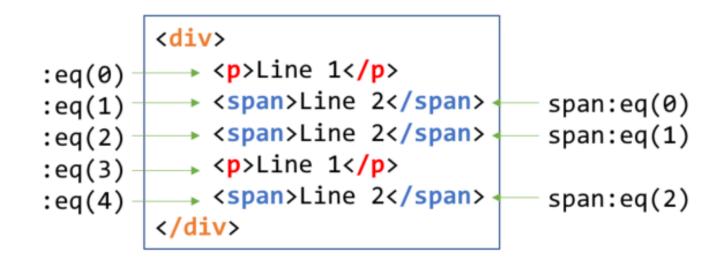
結合子

名称	構文	説明	例
子孫結合子	АВ	子孫にあたるノード	#MainContent p
子結合子	A > B	直接の子のノード	#MainContent > div > p
一般兄弟結合子	A ~ B	Aの後にあるBの要素	p ~ p
隣接兄弟結合子	A + B	Aの直後のBの要素	span + p

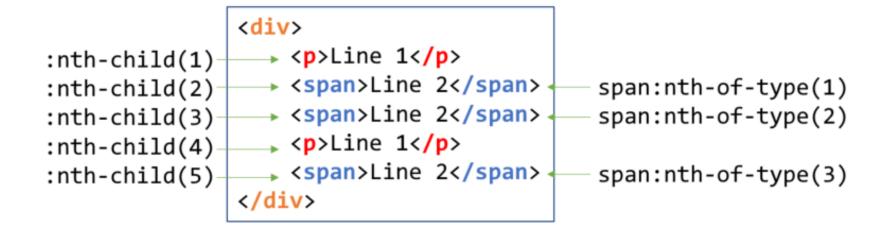


フィルター

構文	説明
:first	最初の要素
:last	最後の要素
:even	偶数番目の要素
:odd	奇数番目の要素
:eq()	指定したインデックス番号の要素
:lt()	指定したインデックスより小さい番号の要素
:gt()	指定したインデックスより大きい番号の要素
:not()	動作未確認



構文	説明
:first-child	最初の子要素
:first-of-type	セレクタで指定された 最初の子要素
:last-child	最後の子要素
:last-of-type	セレクタで指定された 最後の子要素
:only-child	子要素が 1つだけのときに取得
:only-of-type	セレクタで指定された要素が 1つだけのときに取得
:nth-child()	親要素内にあるセレクタで 引数で指定した要素
:nth-of-type()	セレクタと引数で指定された要素
:nth-last-child()	nth-childを最後から数える
:nth-last-of-type()	nth-of-typeを最後から数える

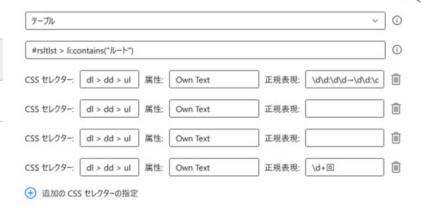


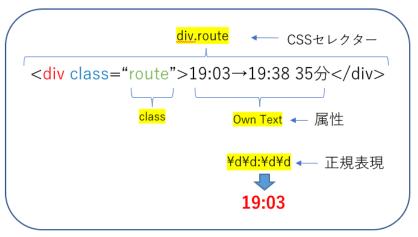


:eq()、:lt()、:gt()は、0から始まります。

:nth-child()、:nth-of-type():nth-last-child()、:nth-last-of-type()は 1から始まります

構文	説明
:contains()	引数で指定した文字列を含む要素
:has()	指定したセレクタを「子孫要素」に持つ要素
:empty	Own Textが何もない要素
:parent	動作未確認





CSSセレクター

- 1. dl > dd > ul > li
- 2. dl > dd > ul > li > span:last-of-type
- 3. dl > dd > ul > li:eq(1)
- 4. dl > dd > ul > li:eq(2)

正規表現

- 1. ¥d¥d:¥d¥d→¥d¥d:¥d¥d
- 2.
- 3.
- 4. ¥d+回

Yahoo!路線情報のデータ取得

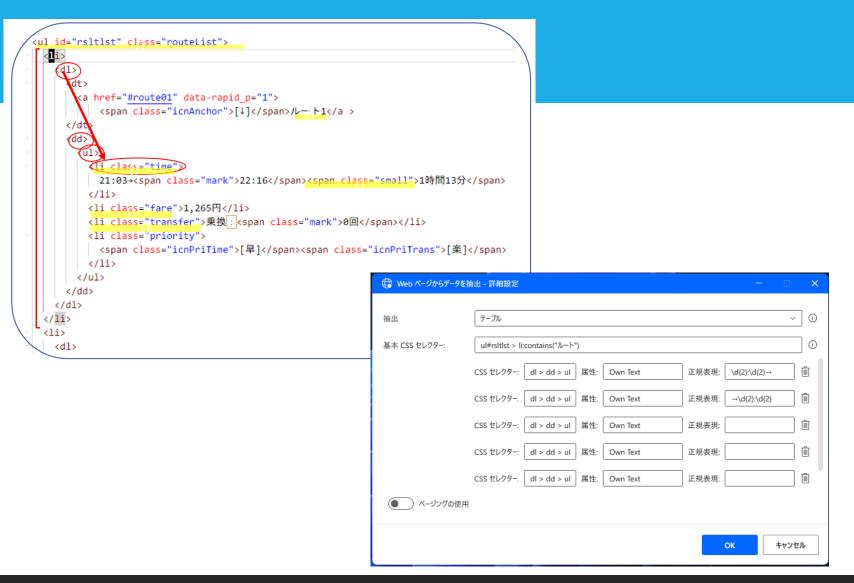
9 3	到着時刻順	樂乗換回]数順	安料	金の安い順
↓ルート1	19:12→ <mark>20:30</mark>	1時間18分	1,265円 美	乗換:0回	<i>9</i>
↓ルート2	19:07→ <mark>20:30</mark>	1時間23分	1,265円 ∌	乗換:1回	9
√ルート3	19:23→20:36	1時間13分	1,265円 美	乗換:0回	%

変数の値

DataFromWebPage (Datatable)

#	Value #1	Value #2	Value #3	Value #4	Value #5
0	19:55→	→21:12	1時間17分	1,265円	乗換:0回
1	19:52→	→21:12	1時間20分	1,265円	乗換:1回
2	19:58→	→21:26	1時間28分	1,265円	乗換:0回

Yahoo!路線情報のデータ取得



基本CSSセレクター

ul#rsltlst > li:contains("ルート")

CSSセレクター

- 1. dl > dd > ul > li.time
- 2. dl > dd > ul > li.time
- 3. dl > dd > ul > li.time > span.small
- 4. dl > dd > ul > li.fare
- 5. dl > dd > ul > li.transfer

正規表現

- 1. \d{2}:\d{2}→
- 2. →\d{2}:\d{2}

Power Automate for desktop CSSセレクタで Web操作が快適に

おわり