

Web Scraping

Automation 360 CE

Prof. Osvaldo Kotaro Takai Profa. Ana Cristina dos Santos

Objetivos

Acompanhe, a seguir, os objetivos de aprendizagem para esta parte:

Aprender a extrair dados de páginas web.

O que vamos fazer?

- 1. Criaremos um Bot que:
 - 1.1. Navega para a página web: <u>Aprovadores do</u> Fundão.
 - 1.2. Extrai:
 - 1.2.1. O Título da matéria.
 - 1.2.2. A tabela dos deputados que votaram a favor do afrouxamento das regras eleitorais contendo três colunas: NOME, PARTIDO e ESTADO.
 - 1.3. Salva essa tabela numa arquivo CSV com o nome do título da matéria.

Figura 1 - Esquema do Bot.







11 - Obter dados da Web





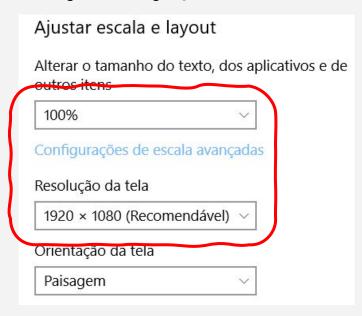
Quem foram os 263 deputados que alteraram (pra pior) as regras eleitorais

Fonte: do autor, 2021.

Preparação - Configuração do Windows para a Gravação de Scripts

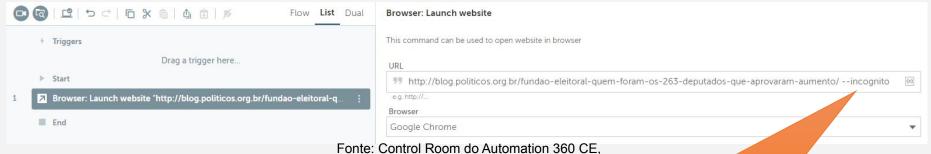
- → Vá para Configurações do Windows / Sistema / Vídeo.
- → Altere o tamanho do texto, dos aplicativos e de outros itens para 100%.
- → Coloque a Resolução da tela no Recomendável.

Figura 2 - Configuração no word.



Fonte: do autor, 2021.

Figura 3 - Configuração da ação Browser: Launch website.



2021.

Figura 4 - Página web aberta com a opção "--incognito".

Depois de configurar, salve e execute.



Fonte: Blog Ranking dos Políticos, 2021.

http://blog.politicos.org.br/fundao-ele itoral-quem-foram-os-263-deputadosque-aprovaram-aumento/

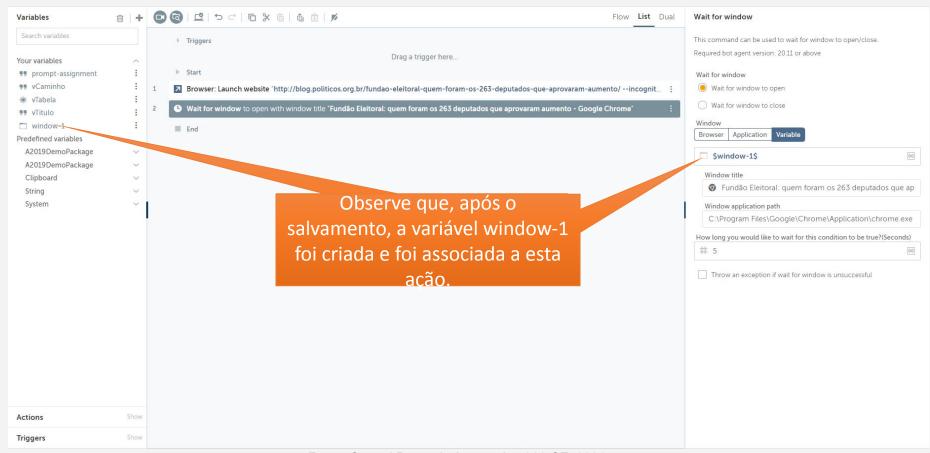
Ao final, coloque o parâmetro "--incognito" sem acento para que o navegador **chrome** seja aberto no modo Anônimo.

Figura 5 - Configuração da ação "Wait for window".



Figura 6 - Configuração da ação "Wait for window". Flow List Dual Wait for window † Triggers This command can be used to wait for window to open/close Required bot agent version: 20.11 or above Drag a trigger here.. ▶ Start Wait for window Browser: Launch website "http://blog.politicos.org.br/fundao-eleitoral-quem-foram-os-263-deputados-que-aprovaram-aumento/ --incognit... ; Wait for window MA : **▲** Window End Browser Application Variable Choose an application window Selecione Application; Presets Currently active active when the bot is running Atualize o combobox; e Desktop default space behind open windows 3) Selecione o navegador Taskbar shows open applications Available windows Chrome com a página teste | Edit Task Bot | Control Room | Automation . Fundão Eleitoral: quem foram os 263 deputados q... anteriormente aberta. Unidade 3 - Parte 3 - Scalping - Apresentações Go... Sem título* - Paint 3D Q-Dir 9.81 Salve Monitor OLEChannel Wnd Serviços

Figura 7 - Variável window-1 criada.



Fundão Eleitoral: quem foram os 🗙 A Não seguro | blog.politicos.org.br/fundao-eleitoral-quem-foram-os-263-deputados-que-aprovaram-aumento, Google Tradutor G Google RPA A Compare plataform... Respostas DESTAQUES NOTÍCIAS RANKING O QUE VOCÊ PROCURA? Quem foram os 263 deputados que alteraram (pra pior) as SEARCH Search regras eleitorais 4 SETEMBRO 2019 | RANKING DOS POLÍTICOS LEIA MAIS A Câmara dos Deputados aprovou, por 263 a 144, projeto de lei que abre caminho para ampliar o bilionário fundo eleitoral, recria a propaganda partidária na TV (que foi extinta justamente para custear as campanhas políticas), Indicação de Kássio Nunes ao STF: análise da autoriza o uso de dinheiro público para pagar advogados de investigados e institui regras mais flexíveis para repercussão contratação de dirigentes partidários. Celso Russomanno: como foi o mandato do polêmico deputado federal Por estado: os 42 senadores que votaram para permitir Confira abaixo a lista com os nomes dos deputados que votaram A FAVOR e CONTRA o projeto. Veja também a lista aumentos para o funcionalismo durante a crise organizada por estados, e confira os nomes e votos dos parlamentares do seu estado. Confira quem são os 125 parlamentares que querem concorrer às eleições municipais; maioria vai mal no Ranking dos Políticos Deputados que votaram A FAVOR do afrouxamento das regras eleitorais Confira o Ranking dos estados de acordo com a pontuação de seus parlamentares NOME PARTIDO **ESTADO** Alan Rick DEM AC Jéssica Sales MDB AC Perpétua Almeida PCdoB AC Manuel Marcos PRB Selecione esse texto e dê Dra, Vanda Milani SOLIDARIEDADE AC Isnaldo Bulhões Jr. MDB AL Ctrl+C para copiar. AL Sergio Toledo Arthur Lira AL Paulão AL PT Nivaldo Albuquerque PTB AL Marcelo Ramos

Figura 8 - Página web usada para extrair dados.

Fonte: Blog Ranking dos Políticos, 2021.

Figura 9 - Configuração da ação Simulate keystrokes..



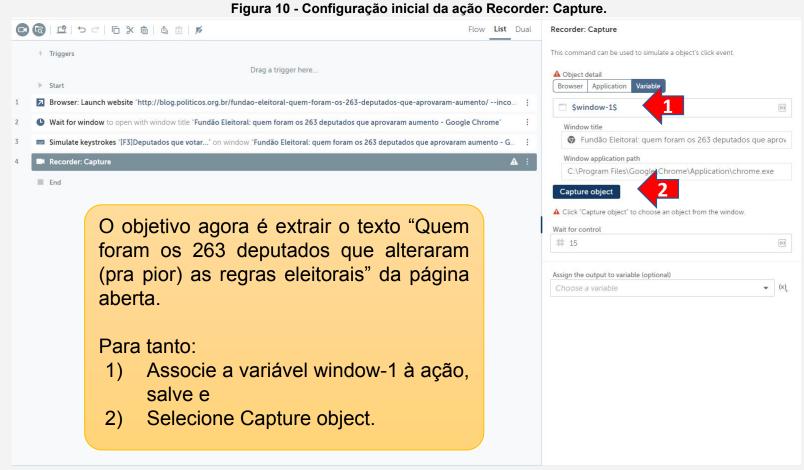


Figura 11 - Seleção do título da matéria.

Quem foram os 263 deputados que alteraram (pra pior) as regras eleitorais DESTAQUES NOTÍCIAS RANKING O QUE VOCÊ PROCURA? Quem foram os 263 deputados que alteraram (pra pior) as Search SEARCH regras eleitorais 4 SETEMBRO 2019 | RANKING DOS I LEIA MAIS A Câmara dos Deputados aprovoú, po 144, projeto de lei que abre caminho para ampliar o bilionário fundo eleitoral, recria a propaganda partidária na TV (que foi extinta justamente para custear as campanhas políticas), Indicação de Kássio Nunes ao STF: análise da autoriza o uso de dinheiro público para pagar advogados de investigados e institui regras mais flexíveis para repercussão contratação de dirigentes partidários. Movimente o ponteiro do mouse sobre a Confira abaixo a lista com os nomes dos deputados que votaram A l página até que itens da página comecem a organizada por estados, e confira os nomes e votos dos parlamenta ser contornados em vermelho. Deputados que votaram A FAVOR do afrouxamento das regras elei 2) Posicione o ponteiro do mouse sobre o texto NOME PARTIDO **ESTADO** que queremos selecionar até que seja Alan Rick DEM. AC completamente contornado por uma linha Jéssica Sales vermelha e dê um clique. Perpétua Almeida PCdoB Manuel Marcos PRB AC Dra. Vanda Milani SOLIDARIEDADE AC Isnaldo Bulhões Jr. MDB AL Sergio Toledo

Fonte: Blog Ranking dos Políticos, 2021.

Figura 12 - Configuração da ação Recorder: Capture após a captura do Título.

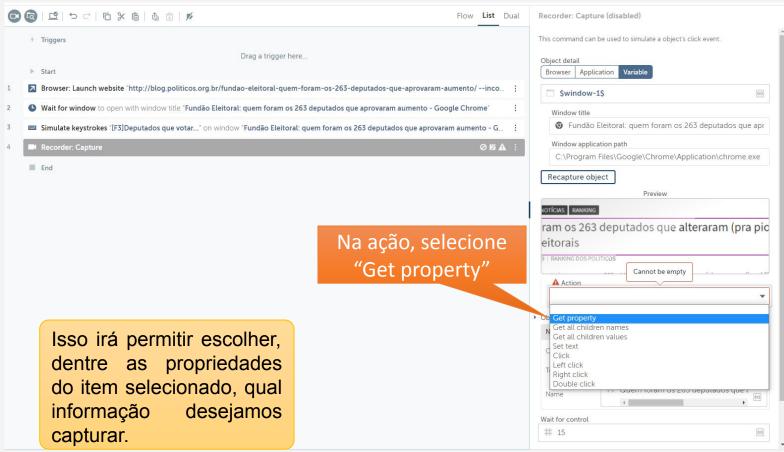
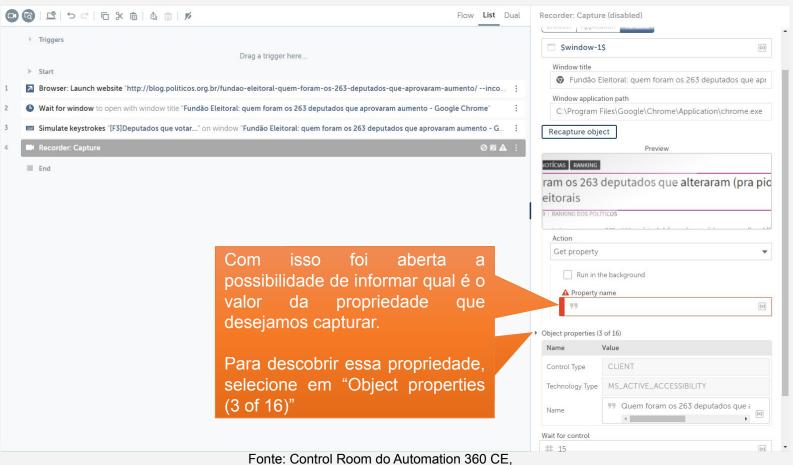


Figura 13 - Configuração da ação Recorder: Capture após abrir a opção Object properties.



2021

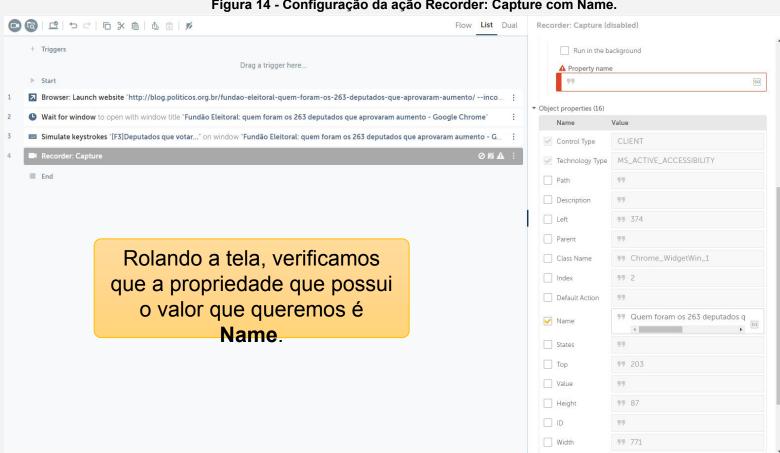


Figura 14 - Configuração da ação Recorder: Capture com Name.

Figura 15 - Configuração da ação Recorder: Capture com Name.

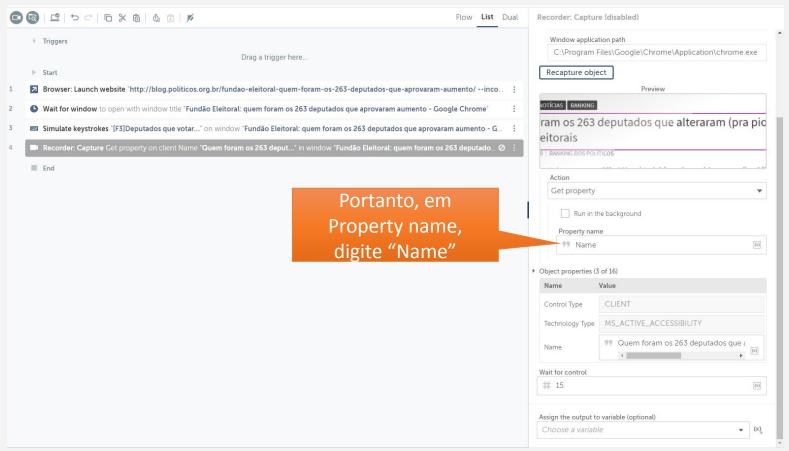




Figura 17 - Configuração da ação Recorder: Capture para capturar tabela.

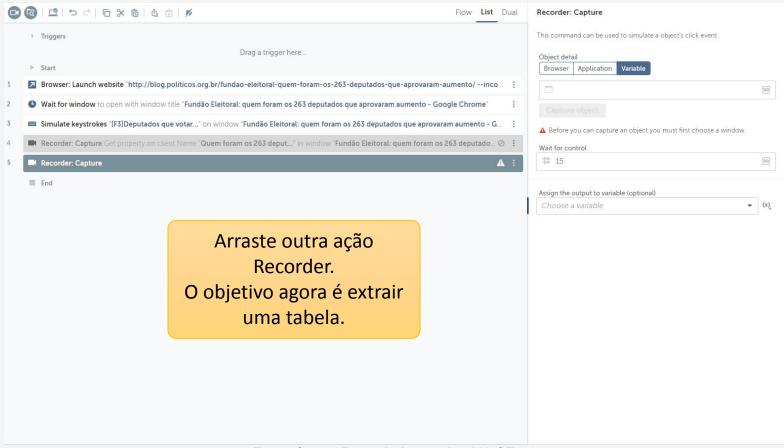


Figura 18 - Configuração da ação Recorder: Capture para capturar tabela.

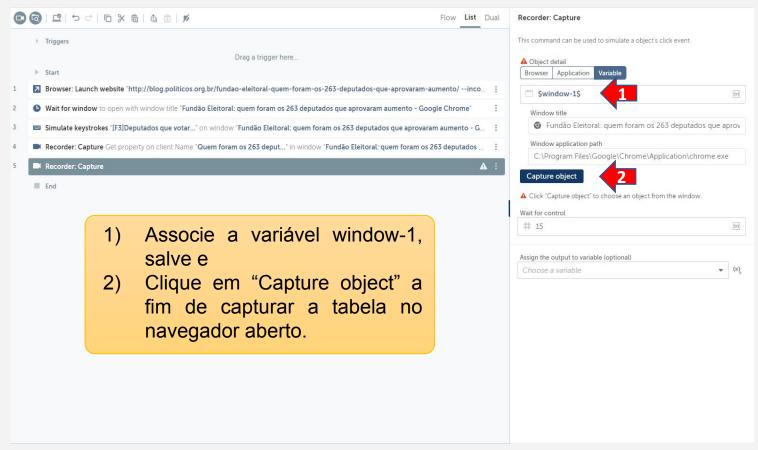


Figura 19 - Página do Chrome aberta para seleção da tabela.

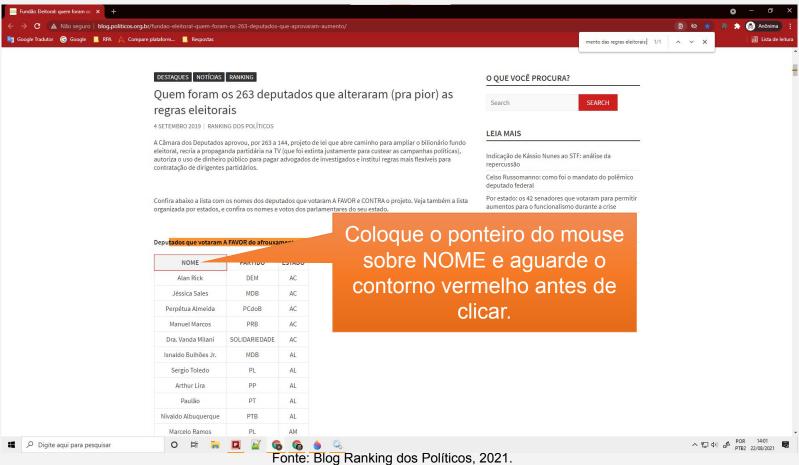
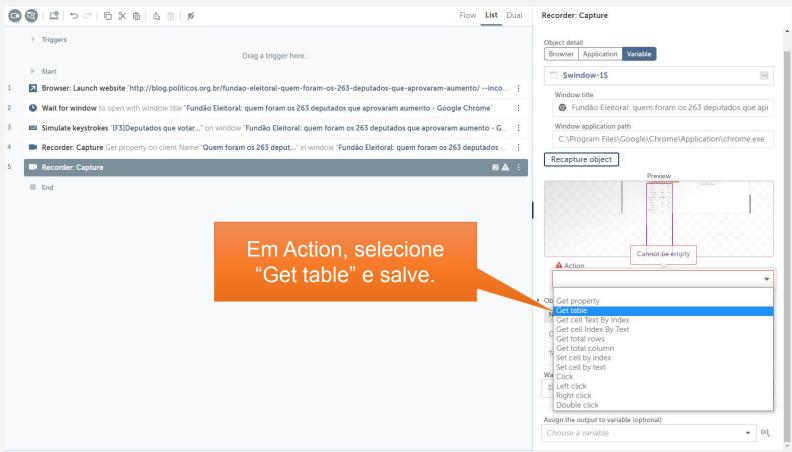
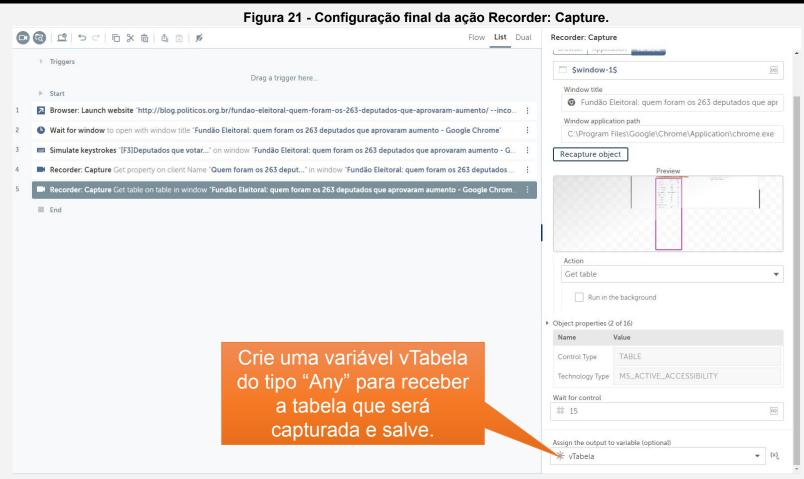


Figura 20 - Configuração inicial da ação Recorder: Capture após capturar a tabela.





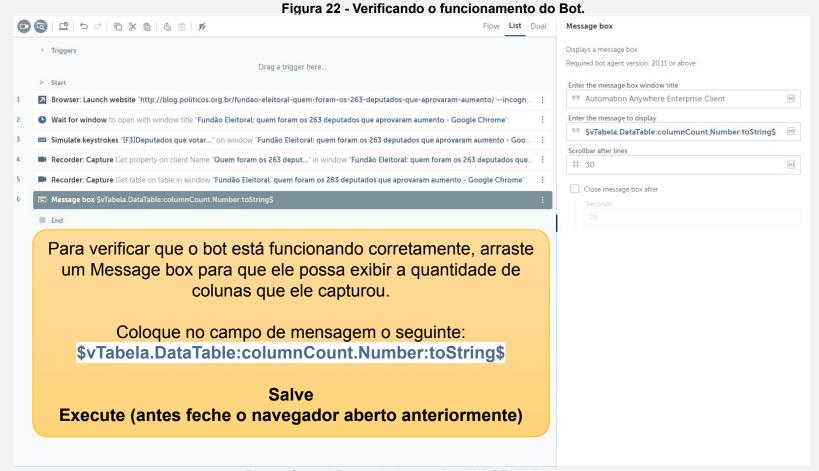
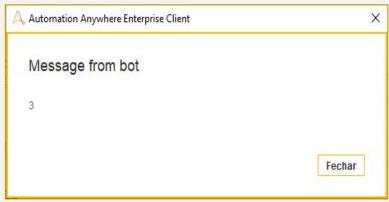


Figura 23 - Mensagem exibida



Fonte: Control Room do Automation 360 CE, 2021.

O resultado deve ser 3. Se isso não ocorreu, garanta que todos os passos anteriores foram seguidos com precisão.

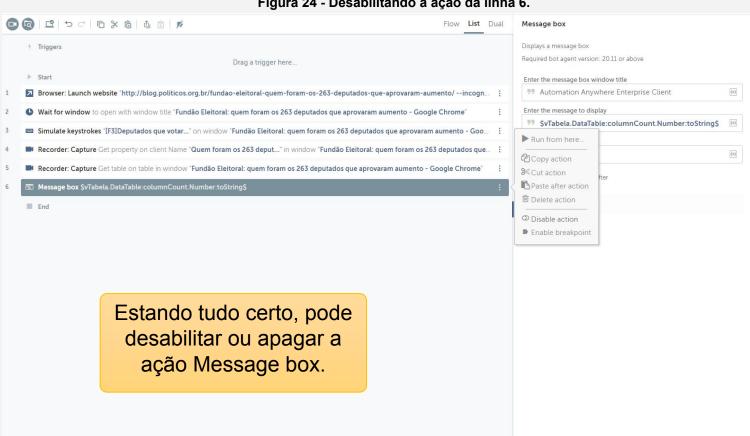
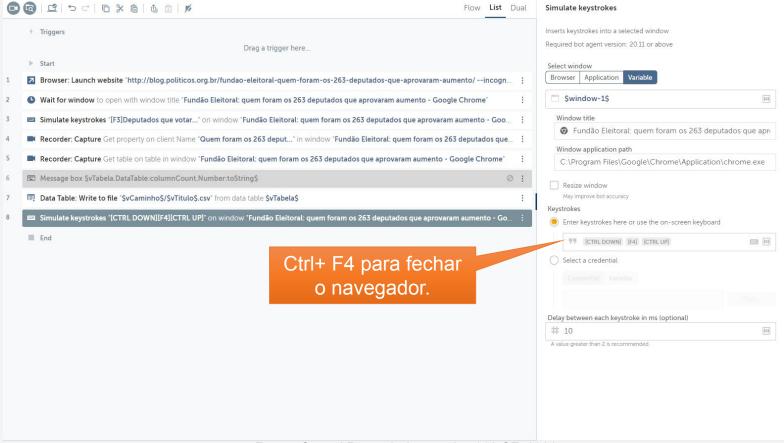


Figura 24 - Desabilitando a ação da linha 6.

Figura 25 - Configuração da ação Data Table: Write to file. Flow List Dual Data Table: Write to file † Triggers Writes a data table into a specified file Required bot agent version: 20.11 or above Drag a trigger here.. ▶ Start Data table name * vTabela ▼ (x) Browser: Launch website "http://bloq.politicos.org.br/fundao-eleitoral-quem-foram-os-263-deputados-que-aprovaram-aumento/ --incor Enter file name Wait for window to open with window title "Fundão Eleitoral: quem foram os 263 deputados que aprovaram aumento - Google Chrome" 99 SvCaminho\$/\$vTitulo\$.csv Browse.. 📼 Simulate keystrokes "[F3]Deputados que votar..." on window "Fundão Eleitoral: quem foram os 263 deputados que aprovaram aumento Required extensions: ".csv", ".txt" 💌 Recorder: Capture Get property on client Name "Quem foram os 263 deput..." in window "Fundão Eleitoral: quem foram os 263 deputados qu Create folders/files if it doesn't exist Recorder: Capture Get table on table in window "Fundão Eleitoral: quem foram os 263 deputados que aprovaram aumento - Google Chrome" Append to the existing file Message box \$vTabela.DataTable:columnCount.Number:toString\$ Overwrite existing file Data Table: Write to file "\$vCaminho\$/\$vTitulo\$.csv" from data table \$vTabela\$ Row delimiter III End new line Column delimiter O objetivo agora é salvar a tabela Encoding UTF-8 capturada num arquivo csv. Para tanto, configura esta ação conforme indicado pelas setas vermelhas.

Figura 26 - Configuração da ação Simulate keystrokes (para fechar o chrome).



Salve e Execute

• Dúvidas?

Exercício

- Faça com que um Bot extraia a tabela dos Deputados que votaram CONTRA o afrouxamento das regras eleitorais e salve-o num outro arquivo .CSV.
- Depois, junte os dois arquivos num único arquivo com as colunas NOME, PARTIDO, ESTADO e COMO-VOTOU (Bem ou Mal).

Deputados que votaram CONTRA o afrouxamento das regras eleitorais:

NOME	PARTIDO	ESTADO
Mara Rocha	PSDB	AC
Jhc	PSB	AL
Sidney Leite	PSD	AM
Capitão Alberto Neto	PRB	AM
Leda Sadala	AVANTE	AP
Tito	AVANTE	ВА
Uldurico Junior	PROS	ВА
Charles Fernandes	PSD	ВА
Idilvan Alencar	PDT	CE
Capitão Wagner	PROS	CE

FIM

