

# *Scraping Certified Inspectors Data From nachi.org by Chedy Chaaben*

There is an estimated total of 11,100 rows of data that needs to be scraped from a directory found at <https://www.nachi.org/certified-inspectors/browse/us/>, and produced in .csv format.

The scraped data shall fill the following columns within each row produced:

FirstName, LastName, CompanyName, City, State, Website, InterNACHI, OfficePhone, MobilePhone, FaxPhone

If data for any of the above fields is not present within the directory listing and respective directory listing pages being scraped, I will put “null” in its place.

Here is an example to reference, and how to identify the data needed:

The screenshot shows a web browser displaying the nachi.org/certified-inspectors/browse/us/ page. The header is blue with the International Association of Certified Home Inspectors logo. Below the header, there's a search bar and a sidebar with options like 'Start Over / Browse All' and 'Certified Home Inspector'. The main content area is titled 'Certified Home Inspectors in United States'. It shows a table with one row of data. The table has columns for 'Member', 'Servicing', and 'Phone'. The data row shows 'Richard Walters, HI-4139' under Member, 'Home Inspections Resource' under Servicing, and '(251) 370-7676' under Phone. There are also edit and delete icons next to the phone number.

Member	Servicing	Phone
Richard Walters, HI-4139 Home Inspections Resource	Theodore	(251) 370-7676

## **FirstName**

- The first blue word in the directory listing
- In this example, it is “Richard”

## **LastName**

- The blue word after the FirstName
- Ignore the comma, and space immediately after the FirstName
- In this example, it is “Walters”

## **CompanyName**

- The grey word found under FirstName and LastName
- In this example, it is “Home Inspections Resource”

## **City**

- The word under the ‘Servicing’ column in the directory listing
- In this example, it is “Theodore”

## **Office Phone**

- The number under the ‘Phone’ column in the directory listing
- In this example, it is “(251) 370-7676”

The remainder of the data must be scraped from the respective directory listing’s directory listing page.

Below, is an example of the directory listing page for the directory listing example used above:

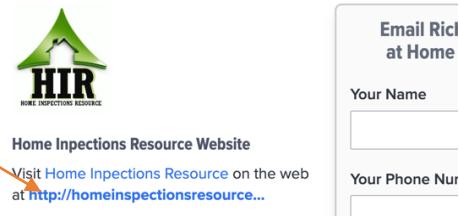
The screenshot shows a web browser window with the URL [nachi.org/certified-inspectors/richard-walters-hi-4139-50089](http://nachi.org/certified-inspectors/richard-walters-hi-4139-50089). At the top, there's a blue header bar for the International Association of Certified Home Inspectors (InterNACHI). Below the header, the name "Richard Walters, HI-4139" is prominently displayed in large, bold letters. To the left of the name is a circular seal for "CERTIFIED PROFESSIONAL INSPECTOR" with "InterNACHI" and "CPI" in the center. Below the name, the text "Home Inspections Resource" is visible. A paragraph below states that Richard Walters is a Theodore, AL Certified Professional Inspector® certified by InterNACHI®. A blue link "\$25,000 Honor Guarantee backed by InterNACHI" is present. A large orange button with a phone icon and the text "Call Richard at (251) 370-7676" is centered. To the right of the button is a small green triangle icon. To the right of the triangle is a white button labeled "Email Richard Walters, HI-4139".

#### State

- Found in the first paragraph of text, following the City (which has already been identified)
- Ignore the comma, and space immediately after the City
- In this example, State is "AL" - just the two letters identifying the respective US state

#### Website

- Found just to the left of the contact form
- It should be the full URL that is being linked to
- In this example, Website is "<http://homeinspectionsresource.com/>"



#### InterNACHI

- Found just to the left of the contact form
- The combination of letters and numbers under the InterNACHI ID section.
- In this example, InterNACHI is "NACHI16041715"

InterNACHI ID  
An InterNACHI® ID can be used to verify Richard Walters, HI-4139's certification at any time.  
 [NACHI16041715](#)

Some directory pages will have a Phone Numbers section to the right of the contact form, and some will not. The example we've been using above does not have a Phone Numbers section.

If there is no Phone Numbers section, MobilePhone and Fax will be "null"

## Mobile Phone

- Found when the Phone Numbers section has a phone number labeled as 'Mobile'
- In this example MobilePhone is "(941) 204-4429"
- It is possible for MobilePhone and OfficePhone to be duplicate values

at <http://www.SWIRL.COM>

### Phone Numbers

📞 (941) 204-4429 Office

📞 (941) 204-4429 Mobile

  
Your Err  
  
Your Me

## Fax

- Found when the Phone Numbers section has a phone number labeled as 'Fax'
- In this example Fax is "(775) 853-2430"
- It is possible for Fax and OfficePhone to be duplicate values

at <http://www.WEINNINSPECTIONSERV...>

### Phone Numbers

📞 (775) 853-2430 Office

📞 (775) 853-2430 Fax

  
Your Email  
  
Your Message

## Additional Inspection Services



chedychaaben@gmail.com



chedychaaben



chedychaaben.com