

# Checkbox Data Export Chrome Extension

## Problem Statement

In many scenarios, agents face difficulty extracting data from websites, especially when the data is presented in large tables or lists. Traditional copy-paste methods are inefficient, error-prone, and cannot handle large amounts of data or structured extraction. Users often require the data in a **structured format** like CSV for analysis, reporting, or processing.

## Solution

The **Checkbox Exporter Chrome Extension** solves this problem by allowing users to **select specific rows using checkboxes** (including pseudo-checkboxes with ARIA roles) and export them directly into a **structured CSV file**. It enables:

- Extraction of large amounts of data from any webpage.
- Preservation of row and column structure in the CSV.
- Easy, click-based selection without coding.
- Support for both real checkboxes and modern web application pseudo-checkboxes.

This solution simplifies and automates the process of extracting large datasets into a structured, usable format.

## Overview

The **Checkbox Exporter** is a Chrome extension designed to help you **export selected checkboxes from web pages into a CSV file**. It supports both:

- **Real checkboxes** (`<input type="checkbox">`)
- **ARIA/pseudo checkboxes** (`role="checkbox"` with `aria-checked` attribute)

This tool preserves the column structure of the rows and works on modern web applications with dynamic content.

---

## Features

1. **Blue Floating Export Bubble**
  - Appears at the bottom-right of every page.
  - Click the bubble to export selected checkboxes into a CSV.
2. **Supports Real and ARIA/Pseudo Checkboxes**
  - Detects standard HTML checkboxes.

- Detects pseudo-checkboxes implemented with `role="checkbox"` and `aria-checked="true"`.
  - 3. **Row Container Detection**
    - Maps each selected checkbox to its closest row container (`tr`, `li`, `div.row`, `article`, `section`, etc.).
    - Extracts the text content of each row cell to build the CSV.
  - 4. **Automatic Column Handling**
    - Groups cells based on their horizontal positions.
    - Maintains column order similar to how they appear on the page.
    - Adds headers automatically if detectable; otherwise uses `Column 1`, `Column 2`, ....
  - 5. **Simple CSV Export**
    - Generates a CSV in memory and triggers a download immediately.
    - Supports thousands of rows per page.
  - 6. **Dynamic Content Friendly**
    - Works on pages where rows or checkboxes are loaded dynamically.
    - Automatically detects visible checkboxes at the time of export.
  - 7. **Cross-page Limitations**
    - This version **exports checkboxes on the current page only**.
    - To export multiple pages, navigate to each page, select checkboxes, and export separately.
- 

## How to Use

### Installation

1. Download and extract the provided `checkbox-exporter.zip`.
2. Open Chrome and go to `chrome://extensions/`.
3. Enable **Developer Mode**.
4. Click **Load unpacked** and select the extracted folder.
5. The extension is now active on all websites.

### Exporting Data

1. Navigate to the web page containing the checkboxes you want to export.
2. Select the checkboxes you wish to export.
3. Click the **blue bubble** at the bottom-right corner.
4. A CSV file (`checkbox_export.csv`) will automatically download.

## Notes

- Only **checked checkboxes** are included in the export.
  - Pseudo-checkboxes must have `role="checkbox"` and `aria-checked="true"` to be detected.
  - The extension **preserves row structure and columns** as detected on the page.
  - For very large pages with thousands of rows, exporting may take a few seconds.
- 

## Tips for Effective Data Extraction

1. **Check all visible checkboxes** you want to include before clicking the bubble.
  2. **Scroll through lazy-loaded pages** to ensure all rows you want are rendered in the DOM.
  3. If the page uses **dynamic or virtualized tables**, ensure all rows are visible; otherwise, only visible rows are exported.
  4. The CSV can be opened in Excel, Google Sheets, or any spreadsheet application.
- 

This extension is designed to make **data extraction simple and reliable** for modern web pages without requiring any coding or scraping tools.