Robotics Articles Scraper – Handoff Guide

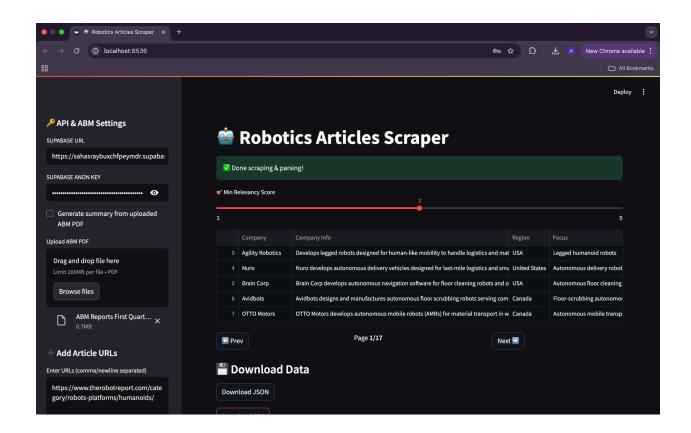
1. Overview

This tool helps non-technical users identify relevant robotics companies from articles online. It extracts and summarizes company details and rates how closely they align with ABM Industries' interests.

2. How to Use the Tool

Step 1: API Setup

- Supabase URL and Anon Key must be entered once (provided by the developer).
- This allows the tool to store and fetch data securely.

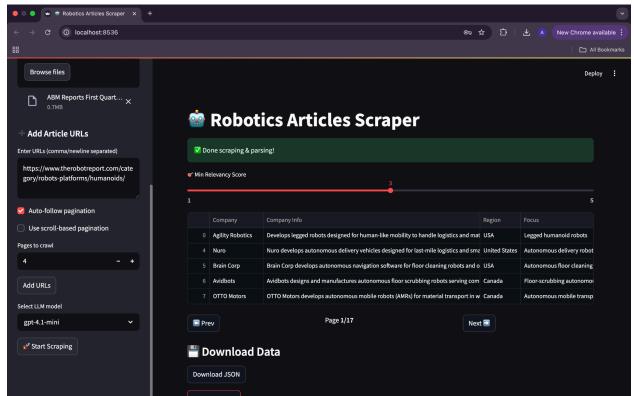


Step 2: Upload ABM PDF

- If you have a quarterly or strategic PDF from ABM, drag and drop it in the "Upload ABM PDF" section.
- This helps tailor company relevance to ABM goals.

Step 3: Add Article URLs

- Paste links from trusted robotics news sources.
 - o Example: https://www.therobotreport.com/category/robots-platforms/hum anoids/
- · Choose either:
 - Auto-follow pagination (recommended)
 - Scroll-based pagination (only for infinite-scroll sites like LinkedIn or Medium)
- Set the number of pages to crawl.



Step 4: Run the Scraper

- Click Start Scraping
- The tool will pull article content, summarize it, and extract company data.

3. Results

Filters

• Use the **Min Relevancy Score** slider to filter companies based on how aligned they are with ABM's needs (scale of 1–5).

Table View

Each row includes:
Company Name, Info, Region, Robotics Focus, Funding, Media Mentions, and more.

Data Export

- After reviewing results:
 - Click **Download JSON** for structured formats.
 - Click **Download CSV** to export spreadsheet-style data.

CSV attached: <u>ScraperFinalOutput.csv</u>

4. What the Relevancy Score Means

Each company is given a **Relevancy Score (1–5)** based on how well it aligns with **ABM Industries' strategic priorities**, such as:

- Smart automation for commercial buildings
- HVAC, janitorial, energy & sustainability solutions
- Robotics that improve facility management
- Commercial-ready solutions (not just research)

If an **ABM PDF** is uploaded, the AI incorporates that context to make even more tailored decisions.

Score	Meaning
5	Strong alignment with ABM's automation/energy goals and facility needs — highly relevant
3	Partial overlap or potential future fit — worth monitoring
1	Low or no relevance — not aligned with ABM's core interests

5. FAQ

Q: Nothing is being scraped?

Make sure you're adding article **pages**, not just category/tag URLs.

Q: CSV is empty?

Try increasing "Pages to crawl" or switching to Auto-follow pagination.

Q: Which model to choose?

Use gpt-4.1-mini for speed + quality balance.