



### 1. Start Point:

We begin from the **LinkedIn public alumni page** or from a **CSV exported manually** via Instant Data Scraper. This avoids direct automated scraping.

### 2. Data Collection:

- *Option 1:* Instant Data Scraper → exports visible alumni details into a CSV.
- *Option 2:* For open websites, n8n can fetch raw HTML using the **HTTP Request Node**.

### 3. Processing in n8n:

- **CSV Import / HTTP Request Node** → bring data into n8n.
- **HTML Extract Node** → extract fields (name, company, role, location).
- **Set Node** → organize & rename fields.
- **Function Node (optional)** → clean, standardize, and remove duplicates.

### 4. Storage & Output:

Data is saved to **Google Sheets (Alumni Master)** or a **CSV/Excel file** for records.

### 5. Final Result:

We get a **clean, de-duplicated alumni database** in tabular format, easy to update regularly and ready for coordinators to use.