**Customer Insights with Qdrant, Python and**

Information Extractor

What does this do?

* **Start Fresh:** The workflow begins by clearing any existing review data in your Qdrant vector store for a specified company, ensuring that you’re working with a clean slate.
* **Scrape TrustPilot Reviews:** It then scrapes the TrustPilot pages for the company to gather recent reviews. Using an HTML extraction node, it pulls out details like review authors, ratings, titles, texts, dates, and URLs.
* **Organize Review Data:** The scraped review data is transformed into a structured list where fields such as review dates (converted to proper date formats), ratings (as numbers), and author details are neatly organized.
* **Store in Qdrant:** Next, the workflow uploads the structured reviews into a Qdrant vector database. During this process, it uses OpenAI’s embedding model to convert the review texts into vector representations for similarity searches.
* **Trigger Analysis Subworkflow:** A subworkflow is triggered with parameters (like company ID and a date range) to prepare for further insights.
* **Cluster Reviews:** A Python code node applies a K-means clustering algorithm on the vector embeddings, grouping similar reviews together. Only clusters with three or more points are kept for analysis.
* **Generate Insights with AI:** An AI agent then reviews the grouped clusters. It summarizes the reviews into a short paragraph, assesses the overall sentiment (e.g., positive, negative, neutral), and suggests potential improvements.
* **Export Results to Sheets:** Finally, the workflow formats the insights and raw review data, then appends these details to a designated Google Sheet, making it easy to track and analyze customer feedback.