

# Intelligent Company Researcher

## About

This is the AI Agent powered by LLM model to research on the company on the different parameters such as About the company, Key place holders, recent news, location, website and business area.

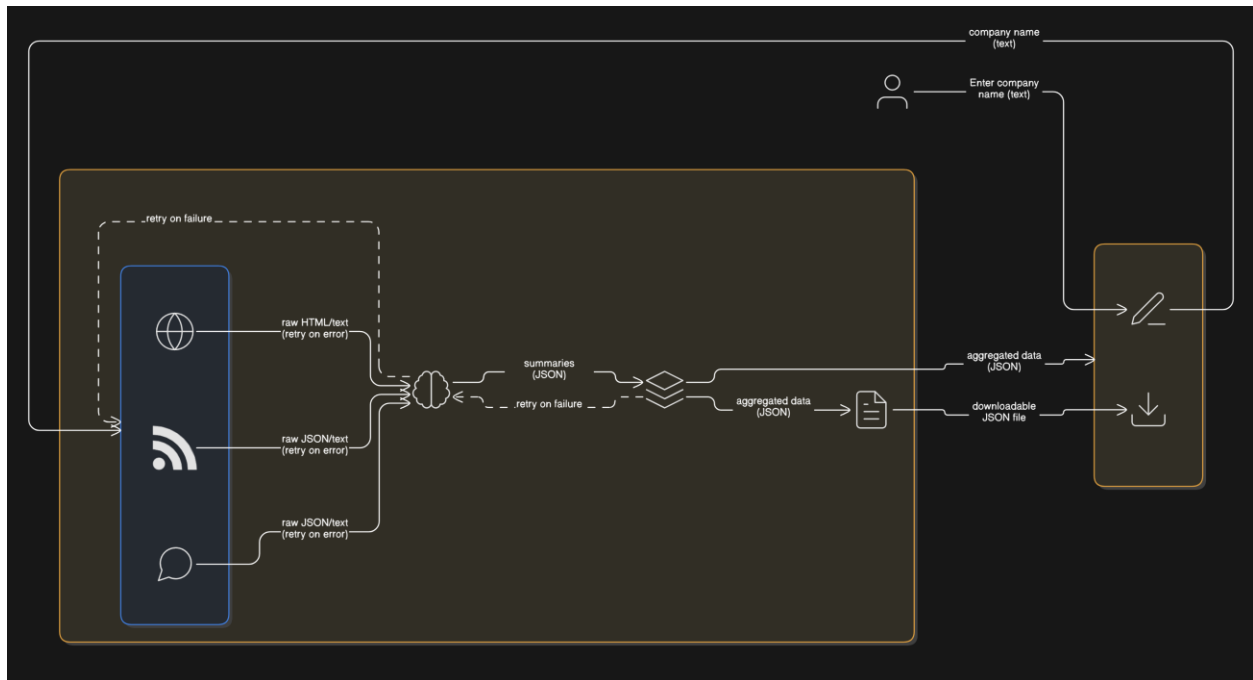
## Architecture Overview

User Input → Company Research Agent

### Company Research Agent:

Module	Function	Output
<code>_get_wikipedia_summary()</code>	Pulls basic summary about the company	Text from Wikipedia
<code>_find_official_website()</code>	Uses DuckDuckGo Instant API to resolve company URL	Official website
<code>_generate_compact_summary()</code>	Uses local LLM (via ollama) to summarize data into 1 paragraph	Natural language summary
<code>_extract_additional_info()</code>	Uses LLM to infer key people, products, and presence	Multiple fact snippets
<code>_get_news_headlines()</code>	Uses Google News RSS feed	List of recent headlines

# Architecture Diagram



## Technology Used

Component	Tool/Technology	Notes
<b>LLM Summarization</b>	Ollama using local models (gemma:2b)	Local lightweight LLM for generating concise company summaries
<b>Search &amp; Info APIs</b>	DuckDuckGo API, Wikipedia REST API	For official site lookup and background company information
<b>News</b>	Google News RSS via feedparser	Used to fetch the latest headlines related to the company
<b>Text Parsing</b>	BeautifulSoup	Included for potential future web scraping needs
<b>HTTP Requests</b>	requests	Manages all API and website communication
<b>Output Format</b>	json module	Stores structured report as downloadable JSON

User Interface	CLI / Streamlit	CLI for local use, Streamlit for interactive web-based input & output
----------------	-----------------	---

# About the LLM Model Used

- Ollama gemma:2b
- Ollama phi

```
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama pull gemma:2b
pulling manifest
pulling c1864a5eb193: 100% 1.7 GB
pulling 097a36493f71: 100% 8.4 KB
pulling 109037bec39c: 100% 136 B
pulling 22a838ceb7fb: 100% 84 B
pulling 887433b89a90: 100% 483 B
verifying sha256 digest
writing manifest
success
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama pull phi
pulling manifest
pulling 04778965089b: 100% 1.6 GB
pulling 7908abcab772: 100% 1.0 KB
pulling 774a15e6f1e5: 100% 77 B
pulling 3188becd6bae: 100% 132 B
pulling 0b8127ddf5ee: 100% 42 B
pulling 4ce4b16d33a3: 100% 555 B
verifying sha256 digest
writing manifest
success
```

- The **ollama gemma:2b** model is used in this system to generate a compact, well-structured summary of the company in a single paragraph. It is a lightweight language model optimized for coherent and fluent natural language generation, making it ideal for summarization tasks where brevity and clarity are essential. Its smaller size ensures faster response times without significantly compromising the quality of text generation, which is particularly useful when generating a general overview of a company that includes its industry, origin, and global presence.
- The **ollama phi model** is integrated into the system to extract specific pieces of information such as key people, products and services, and company headquarters from longer textual data. It is well-suited for instruction-following tasks, where precise and focused responses are required. This model is particularly effective at breaking down unstructured information into clean, actionable outputs, ensuring that targeted company insights are delivered accurately and efficiently. Its ability to follow extraction instructions makes it a strong component for detail-oriented intelligence gathering.

# Ollama Installation

```
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ curl -fsSL https://ollama.ai/install.sh | sh
>>> Installing ollama to /usr/local
[sudo] password for ankitlinux:
>>> Downloading Linux amd64 bundle
##### 100.0%
>>> Creating ollama user...
>>> Adding ollama user to render group...
>>> Adding ollama user to video group...
>>> Adding current user to ollama group...
>>> Creating ollama systemd service...
WARNING: systemd is not running
WARNING: see https://learn.microsoft.com/en-us/windows/wsl/systemd#how-to-enable-systemd to enable it
>>> The Ollama API is now available at 127.0.0.1:11434.
>>> Install complete. Run "ollama" from the command line.
>>> The Ollama API is now available at 127.0.0.1:11434.
>>> Install complete. Run "ollama" from the command line.
```

```
Use "ollama [command] --help" for more information about a command.
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama version
Error: unknown command "version" for "ollama"
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama pull mistral
Error: ollama server not responding - could not connect to ollama server, run 'ollama serve' to start it
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama serve
Couldn't find '/home/ankitlinux/.ollama/id_ed25519'. Generating new private key.
Your new public key is:
```

```
ssh-ed25519 AAAAC3NzaC1lZDI1NTE5AAAAIJBUOGxWR5HH/wIbHiOD1MLkcWJ9txASG61QCN7RzbAg
```

```
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama
```

Usage:

```
ollama [flags]
ollama [command]
```

Available Commands:

serve	Start ollama
create	Create a model from a Modelfile
show	Show information for a model
run	Run a model
stop	Stop a running model
pull	Pull a model from a registry
push	Push a model to a registry
list	List models
ps	List running models
cp	Copy a model
rm	Remove a model
help	Help about any command

Flags:

```
-h, --help      help for ollama
-v, --version   Show version information
```

Use "ollama [command] --help" for more information about a command.

```
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama version
Error: unknown command "version" for "ollama"
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama pull mistral
Error: ollama server not responding - could not connect to ollama server, run 'ollama serve' to start it
ankitlinux@DESKTOP-K1I1BJH:/mnt/c$ ollama serve
Couldn't find '/home/ankitlinux/.ollama/id_ed25519'. Generating new private key.
```

## Sample JSON Output:

```
{  
  "company_name": "google",  
  "website": "https://www.google.com",  
  "summary": "Sure, here's a summary of Google in 3 sentences:\n\nGoogle is an  
American multinational corporation and technology company that focuses on online  
advertising, search engine technology, cloud computing, computer software, and artificial  
intelligence. It is one of the world's most valuable brands and has been referred to as \"the  
most powerful company in the world\" by the BBC.",  
  "key_people": "Sure, here is the list of key people of Google:\n\n* Larry Page\n*  
Sergey Brin",  
  "products_services": "Sure, here's a summary of the main products and services  
offered by Google:\n\n- Online advertising\n- Search engine technology\n- Cloud  
computing\n- Computer software\n- Quantum computing\n- E-commerce\n- Consumer  
electronics\n- Artificial intelligence (AI)",  
  "locations": "The context does not provide any information about the global presence  
or headquarters locations of Google, so I cannot answer this question from the provided  
context.",  
  
  "recent_news": [  
    "Try on styles with AI, jump on great prices and more - The Keyword",  
    "Is AI Killing Google Search? It Might Be Doing the Opposite - WSJ - The Wall Street  
Journal",  
    "Meta Clashes With Apple, Google Over Child Age Check Legislation - Bloomberg"  
  ]  
}
```

## Ollama Server Response:

```
load_tensors: loading model tensors, this can take a while... (mmap = false)
load_tensors: CPU model buffer size = 1594.93 MiB
time=2025-07-25T19:59:42.550+05:30 level=INFO source=server.go:632 msg="waiting for server to become available" status="llm server
t responding"
time=2025-07-25T19:59:42.815+05:30 level=INFO source=server.go:632 msg="waiting for server to become available" status="llm server
adding model"
llama_context: constructing llama_context
llama_context: n_seq_max = 2
llama_context: n_ctx = 8192
llama_context: n_ctx_per_seq = 4096
llama_context: n_batch = 1024
llama_context: n_ubatch = 512
llama_context: causal_attn = 1
llama_context: flash_attn = 0
llama_context: freq_base = 10000.0
llama_context: freq_scale = 1
llama_context: n_ctx_per_seq (4096) < n_ctx_train (8192) -- the full capacity of the model will not be utilized
llama_context: CPU output buffer size = 1.97 MiB
llama_kv_cache_unified: kv_size = 8192, type_k = 'f16', type_v = 'f16', n_layer = 18, can_shift = 1, padding = 32
llama_kv_cache_unified: CPU KV buffer size = 144.00 MiB
llama_kv_cache_unified: KV self size = 144.00 MiB, K (f16): 72.00 MiB, V (f16): 72.00 MiB
llama_context: CPU compute buffer size = 508.25 MiB
llama_context: graph nodes = 637
llama_context: graph splits = 1
time=2025-07-25T20:00:11.331+05:30 level=INFO source=server.go:637 msg="llama runner started in 95.86 seconds"
[GIN] 2025/07/25 - 20:00:29 | 200 | 1m57s | 127.0.0.1 | POST | "/api/generate"
[GIN] 2025/07/25 - 20:00:36 | 200 | 6.571171914s | 127.0.0.1 | POST | "/api/generate"
[GIN] 2025/07/25 - 20:00:44 | 200 | 8.05826033s | 127.0.0.1 | POST | "/api/generate"
[GIN] 2025/07/25 - 20:00:50 | 200 | 5.665684035s | 127.0.0.1 | POST | "/api/generate"
^Cankit@linux@DESKTOP-K11B3H:/mnt/c/Users/LENOVO FLEX$
```

## Demo Videos:

### Streamlit:

[https://drive.google.com/file/d/1oy1NyIS5zRX\\_dwBxSIpntZWF6jtma4OP/view?usp=sharing](https://drive.google.com/file/d/1oy1NyIS5zRX_dwBxSIpntZWF6jtma4OP/view?usp=sharing)

### CLI:

[https://drive.google.com/file/d/1m91GkIz7j47cT\\_a48n7AJxV3FgwH89AS/view?usp=sharing](https://drive.google.com/file/d/1m91GkIz7j47cT_a48n7AJxV3FgwH89AS/view?usp=sharing)