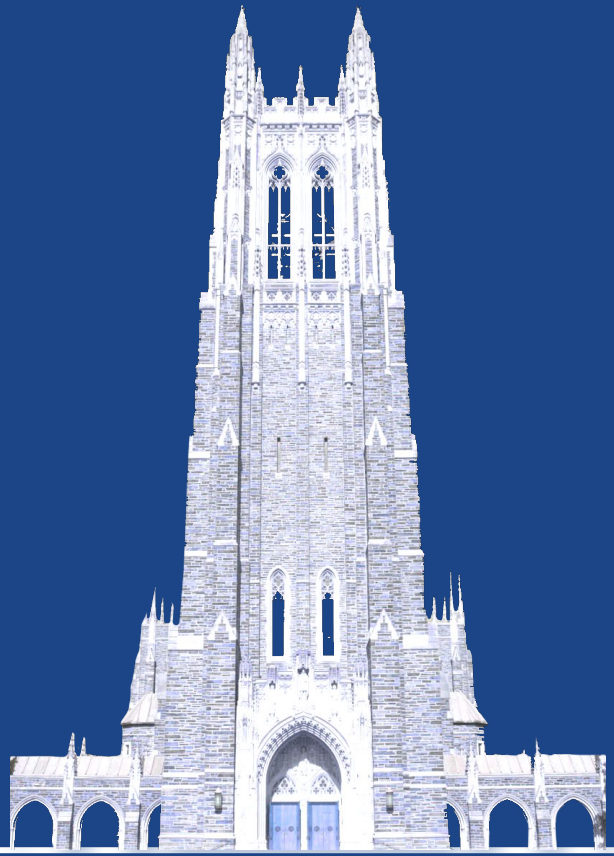
DIIG

SunCast Media

Final Deliverable S24

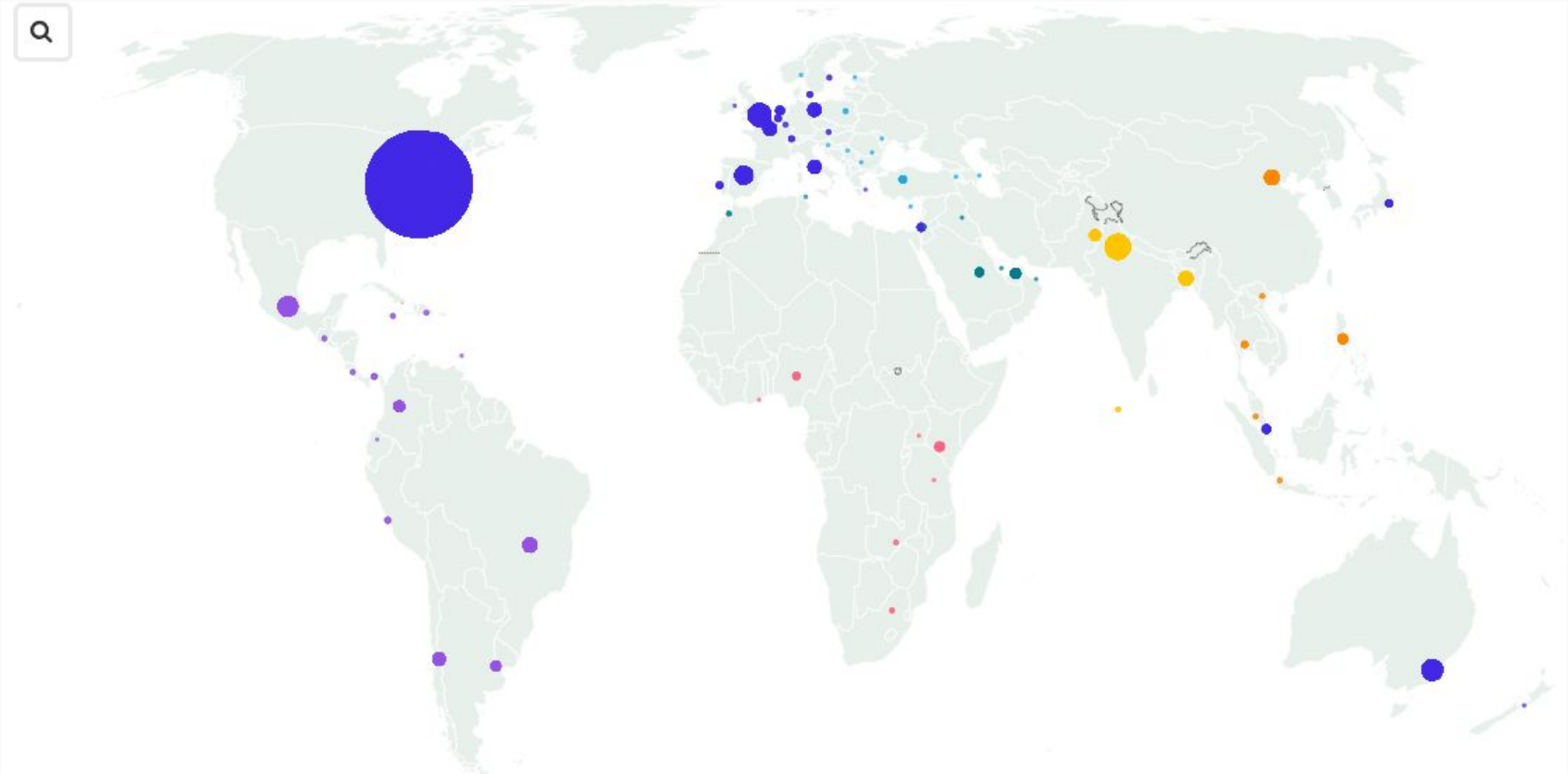
Luke Fu
Ananya Agrawal
Alayna Binder
Rocky Zhang
Yule Fu



1. Audience Analysis
2. Content Analysis
3. Competitor Analysis
4. Follower Classification

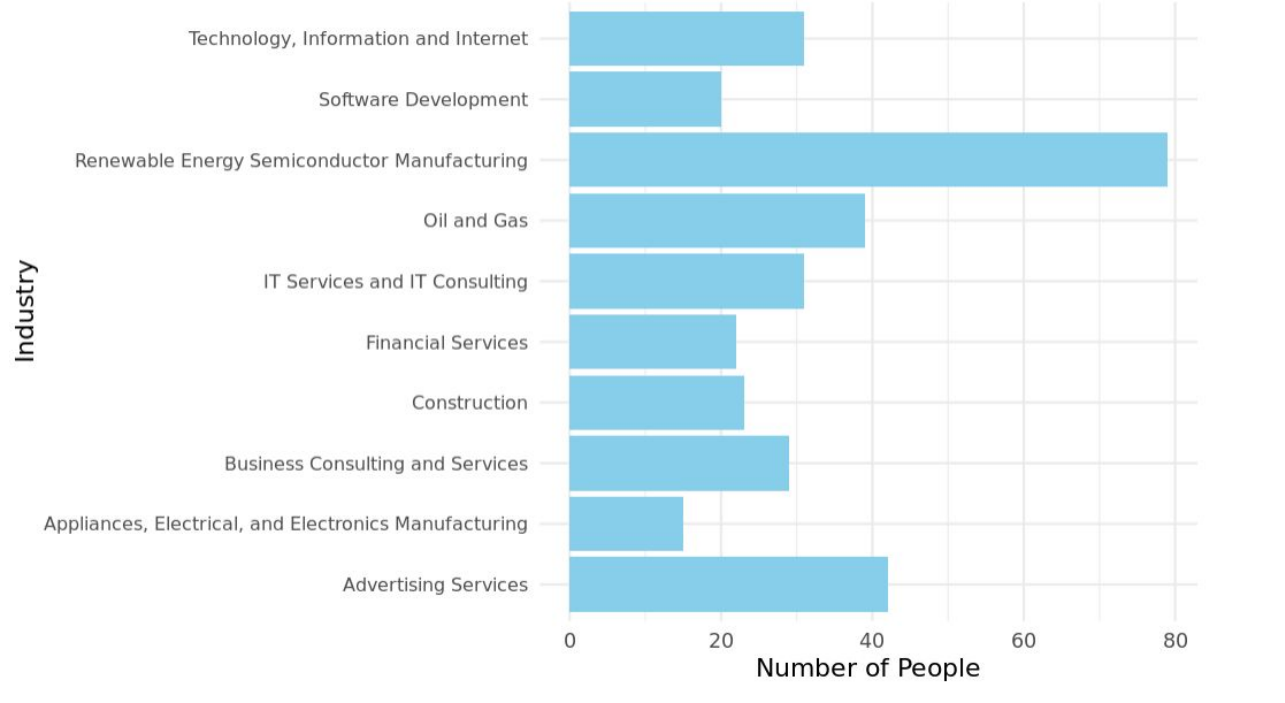
Audience Analysis

Geographical Data



User Data

Given people who aren't in the renewable industry, what industries are they in? Are they open to job switching?



Content Analysis

What type of content performs best?

On YouTube and LinkedIn

⇒ *Evaluate using post performance metrics:* # of views, likes, comments, and/or shares

DIIG

Scraping Process

1. YouTube API
Incorporate YouTube functions into own application.
Retrieved desired information: # of views/likes/comments, video title, publish date

2. Pandas
Data manipulation.
Consolidate scraped data into a dataframe to perform keyword analysis and create visualizations

Keyword Analysis

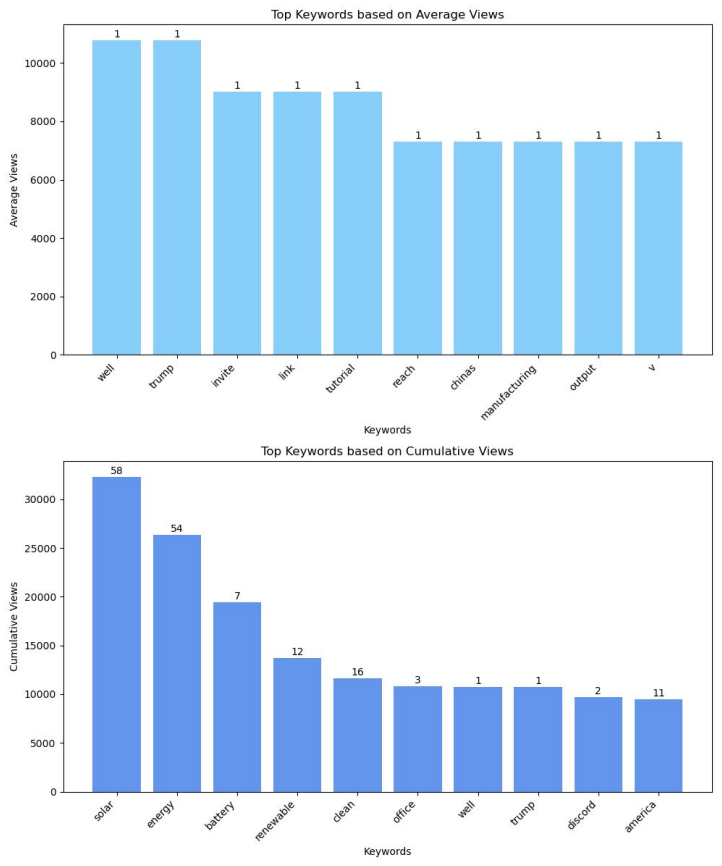
- Idea: Analyze # of views/likes/comments on videos based on title keywords
    - Do specific keywords make videos perform better, on average and overall (cumulatively)?

- Filter to only consider “meaningful” words (i.e., no prepositions)

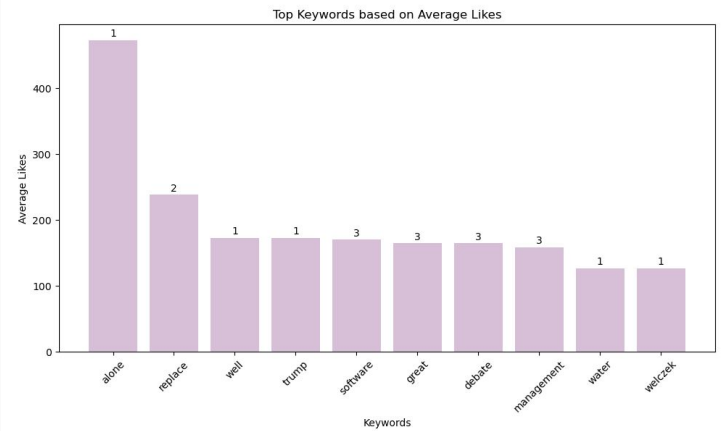
Findings

- Videos on **controversial topics** performed well in average & cumulative # of likes and comments
    - Ex. “Trump” from the title “Renewable Energy Did Well With TRUMP in Office 🔴🇺🇸”
- Videos about “**software**” were well received (high # of likes)
    - Growing sector in energy ⇒ potential topic for further content
- Videos including the keywords “**debate**” and “**illegal**” had among the highest # of comments
    - Ex. “Will Gas Stoves Be Made Illegal?”

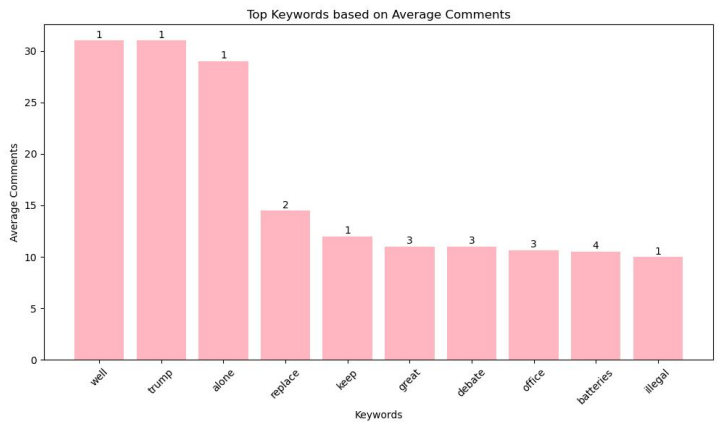
YouTube


Top Keywords based on Average Views
Average Views
Keywords
well
trump
invite
link
tutorial
reach
chinas
manufacturing
output
v
Top Keywords based on Cumulative Views
Cumulative Views
Keywords
solar
energy
battery
renewable
clean
office
well
trump
discord
america

YouTube


Top Keywords based on Average Likes
Average Likes
0
100
200
300
400
1
2
1
1
3
3
3
3
1
1
alone
replace
well
trump
software
great
debate
management
water
welczek
Keywords
Top Keywords based on Cumulative Likes
Cumulative Likes
0
200
400
600
800
1000
1200
60
58
3
4
3
3
2
3
1
11
solar
energy
software
batteries
great
debate
replace
management
alone
ep
Keywords

YouTube


Top Keywords based on Average Comments
Average Comments
Keywords
well
trump
alone
replace
keep
great
debate
office
batteries
illegal
Top Keywords based on Cumulative Comments
Cumulative Comments
Keywords
solar
energy
batteries
renewable
debate
great
office
well
trump
management

LinkedIn

Scraping Process

1. Selenium
Browser automation tool.
Automated opening LinkedIn, logging in, navigating to company page, and scrolling through all posts

2. BeautifulSoup
HTML parser.
While Selenium is scrolling, extract desired features: # of likes/comments/shares, media type, post text, and post date

3. Pandas
Data manipulation.
Consolidate scraped data into a dataframe to perform keyword analysis and create visualizations

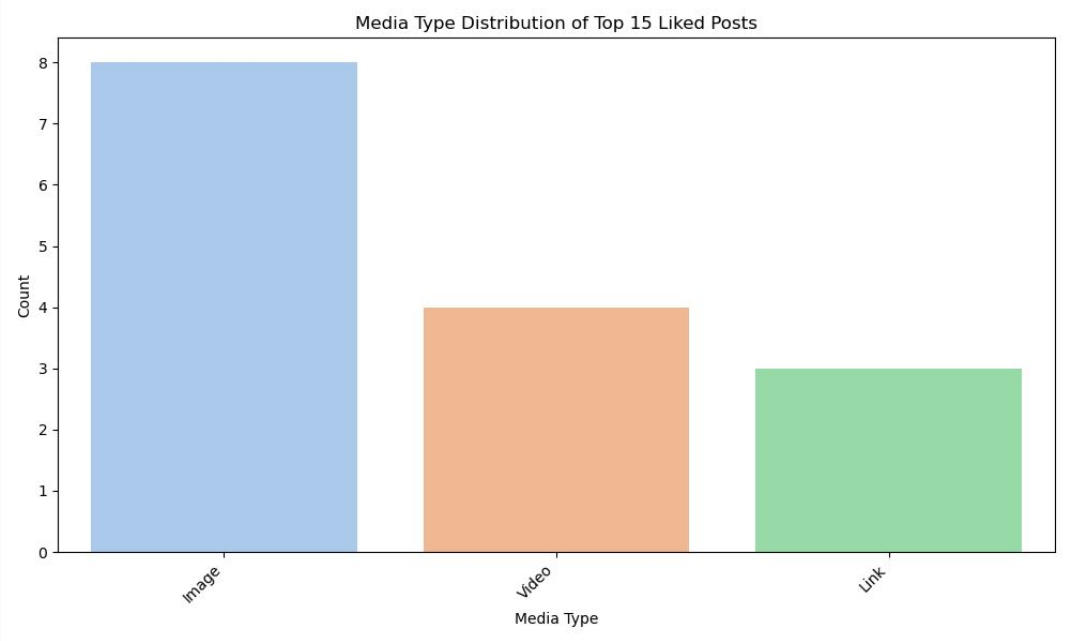
LinkedIn (SunCast)

Findings

- Posts that highlight other **important figures** and **companies** in the clean energy field performed well in # of likes
    - Ex. Martin Pochtaruk (most liked), Daniel French, Nextracker Inc.
- Posts about **oil** received many likes and comments
- Posts about **agrivoltaics** performed well across the board, receiving the highest average # of comments and shares out of all posts
    - Growing sector in energy ⇒ potential topic for further content

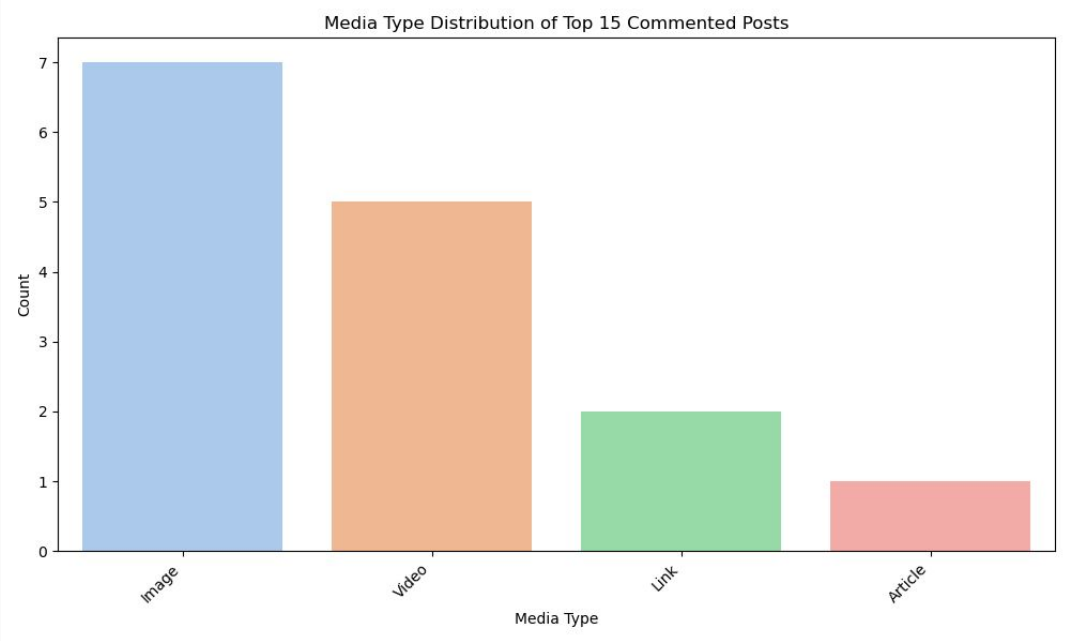
Media Type Analysis

- **Images** performed better in # of **likes**
- Image/Video/Link/Article Ratio
    - **Likes = 8/4/3/0**



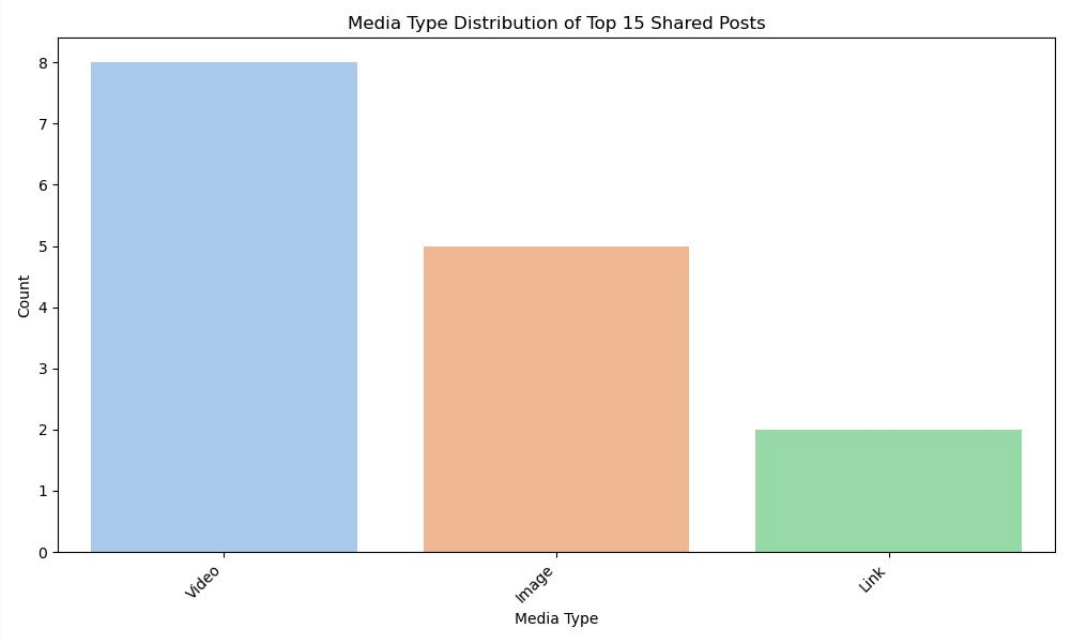
Media Type Analysis

- **Images** performed better in # of **comments**
- Image/Video/Link/Article Ratio
    - **Comments = 7/5/2/1**



Media Type Analysis

- **Videos** performed better in # of **shares**
- Image/Video/Link/Article Ratio
    - **Shares = 5/8/3/0**



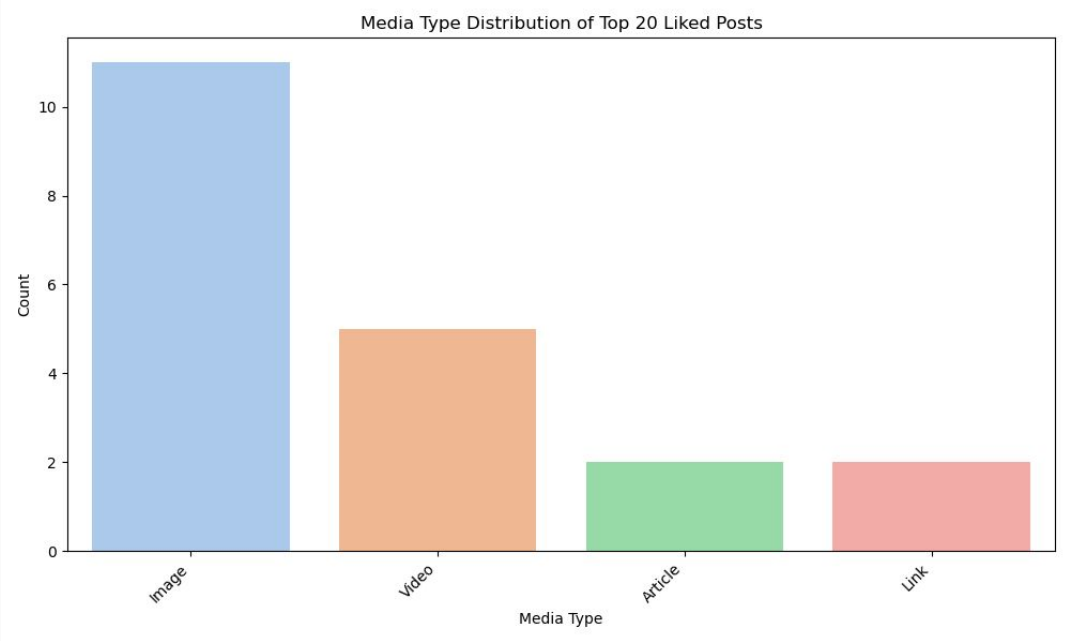
LinkedIn (Nico)

Findings

- Posts at **energy conferences** performed well in # of likes
    - Ex. RE+, GreenExpo (even though 5 years ago, 2nd most liked!)
- What do posts with high # of comments have in common? **Prompting readers**
    - Ex. “What are YOU…?”, “What about you?”, “What trends do you see…?”
- Post about **Houston** received many likes (“clean energy capital of US”)
    - Expand content to global initiatives? Include any breakthrough/unique technologies?
- **Milestones/inspiration** about SunCast well received (100 posts, #becauseofSunCast)

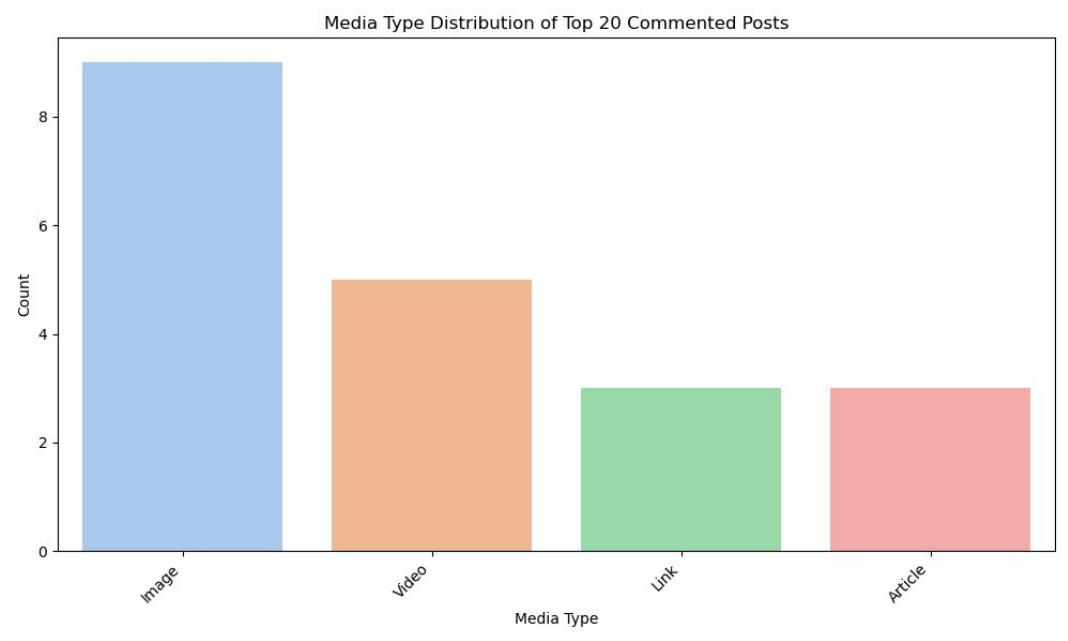
Media Type Analysis

- **Images** performed better in # of **likes**
- Image/Video/Link/Article Ratio
    - **Likes = 11/5/2/2**



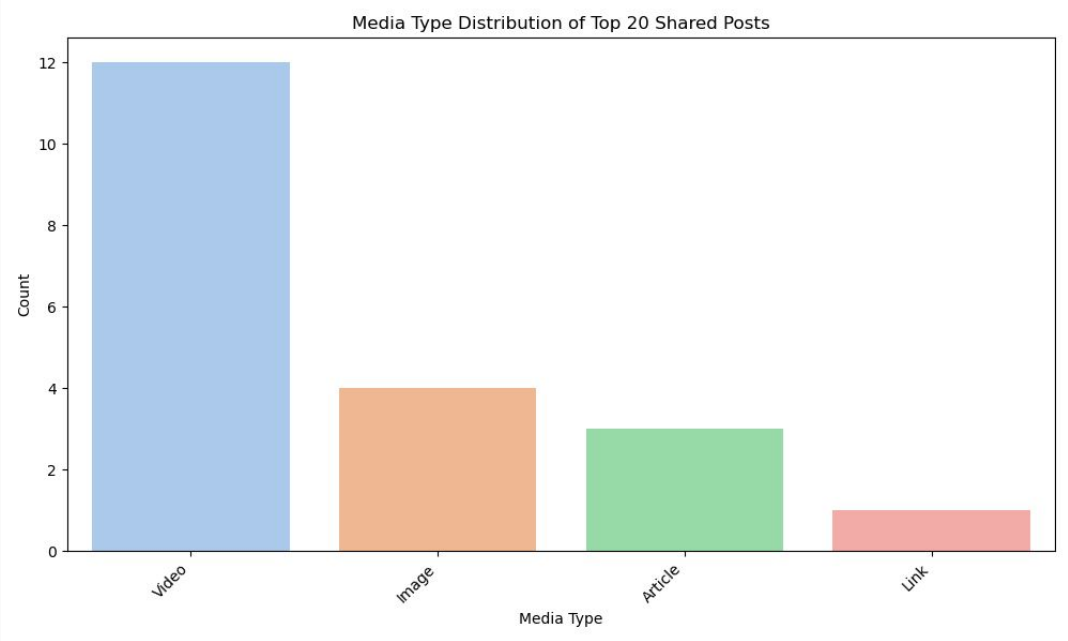
Media Type Analysis

- **Images** performed better in # of **comments**
- Image/Video/Link/Article Ratio
    - **Comments = 9/5/3/3**



Media Type Analysis

- **Videos** performed better in # of **shares**
- Image/Video/Link/Article Ratio
  - **Shares = 12/4/3/1**

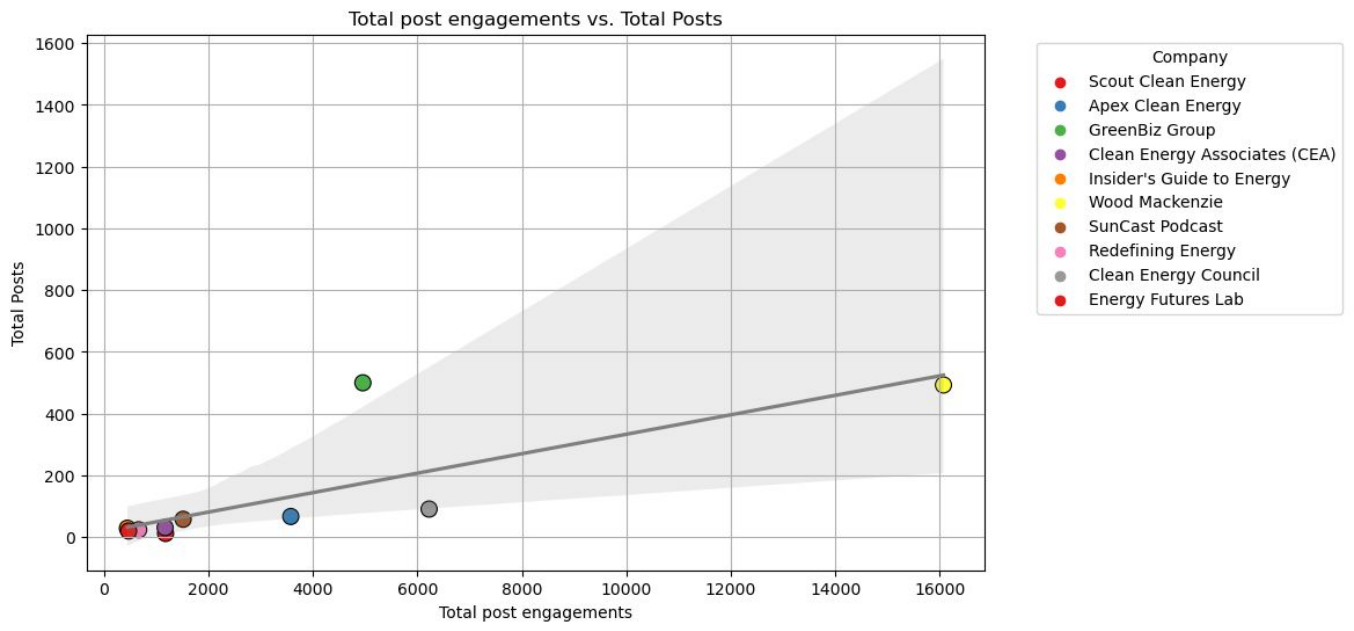


DIIG

Competitor Analysis

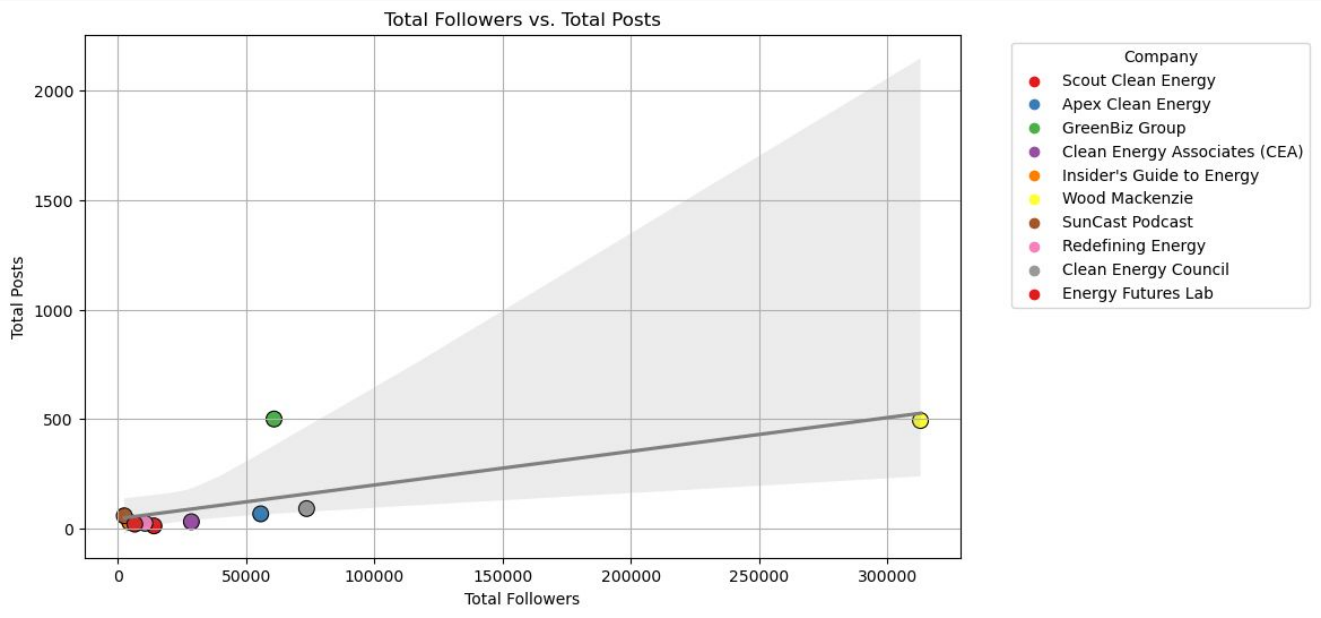
Total post engagements vs. Total Posts

- There is a positive correlation between the number of post engagements (Reactions + Comments + Shares) / Impression) and total posts of a podcast company (90 days).
- Correlation coefficient: **0.785**



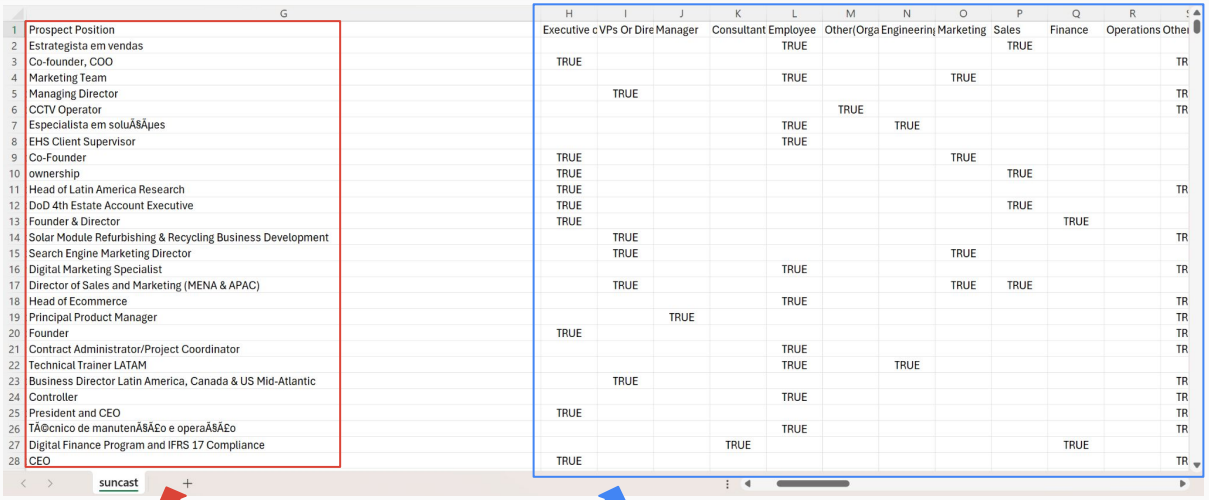
Total Followers vs. Total Posts

- There is a positive correlation between the number of posts and total followers of a podcast company (90 days).
- Correlation coefficient: **0.745**



Follower Classification

Position Classification Model



	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Prospect Position	Executive c	VPs Or Dire	Manager	Consultant	Employee	Other(Orga	Engineerin	Marketing	Sales	Finance	Operations	Other
2	Estrategista em vendas					TRUE				TRUE			
3	Co-founder, COO	TRUE											TR
4	Marketing Team					TRUE			TRUE				
5	Managing Director		TRUE										TR
6	CCTV Operator						TRUE						TR
7	Especialista em soluÃ§Ãµes					TRUE		TRUE					
8	EHS Client Supervisor					TRUE							
9	Co-Founder	TRUE							TRUE				
10	ownership	TRUE								TRUE			
11	Head of Latin America Research	TRUE											TR
12	DoD 4th Estate Account Executive	TRUE								TRUE			
13	Founder & Director	TRUE									TRUE		
14	Solar Module Refurbishing & Recycling Business Development		TRUE										TR
15	Search Engine Marketing Director		TRUE						TRUE				
16	Digital Marketing Specialist					TRUE							TR
17	Director of Sales and Marketing (MENA & APAC)		TRUE						TRUE	TRUE			
18	Head of Ecommerce					TRUE							TR
19	Principal Product Manager			TRUE									TR
20	Founder	TRUE											TR
21	Contract Administrator/Project Coordinator					TRUE							TR
22	Technical Trainer LATAM					TRUE		TRUE					
23	Business Director Latin America, Canada & US Mid-Atlantic		TRUE										TR
24	Controller					TRUE							TR
25	President and CEO	TRUE											TR
26	TÃ©cnico de manutenÃ§Ã£o e operaÃ§Ã£o					TRUE							TR
27	Digital Finance Program and IFRS 17 Compliance				TRUE						TRUE		
28	CEO	TRUE											TR

suncast

written response -> categories

True positive rate: 0.699
True negative rate: 0.939

Fine-tune GPT 3.5

Company Classification Agent - Challenges

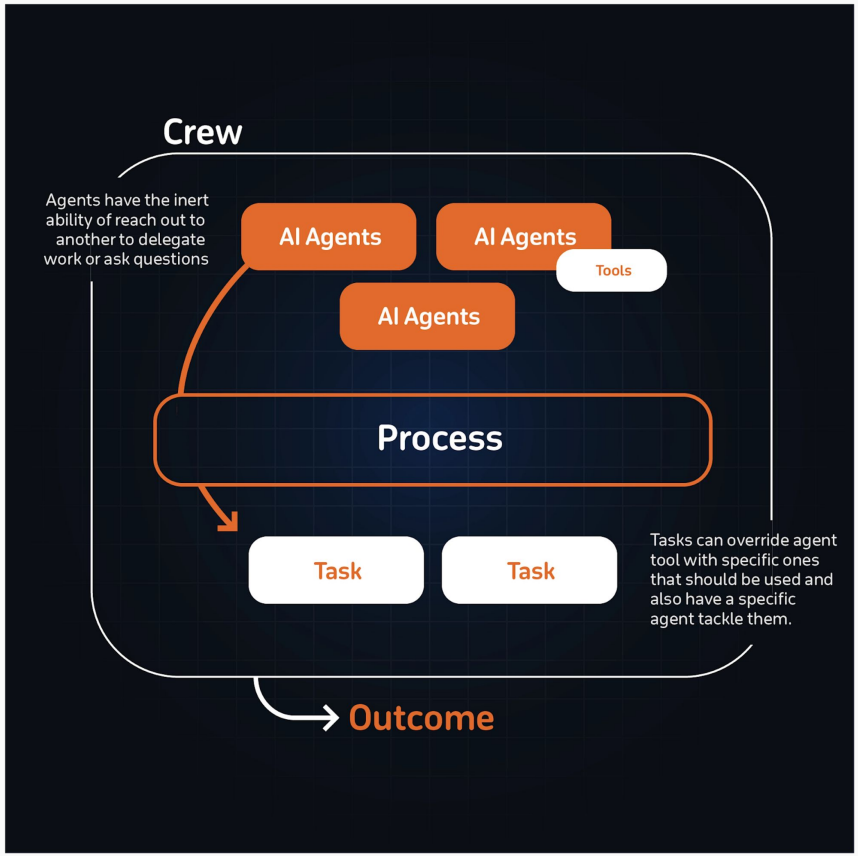
Similar to position classification:
Website url -> categories

Challenges:
Finetune:
- Input not directly classifiable

Scrape then finetune:
- Diverse website structure
- Token limit
- High cost

AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF
Company Website	Company Domain	Solar	Inverter	Energy Storage	Electric Vehicle	Wind	Geothermal	Green Hydroge	Other (Renewat
N/A	N/A								
https://www.taazaa.com/	www.taazaa.com								
http://www.varifarma.com	www.varifarma.com								
http://www.labstart.xyz	www.labstart.xyz								
http://www.q4s.com/complianc	www.q4s.com								
N/A									
https://www.uscompliance.com	www.uscompliance.com								TRUE
www.e1evate.me	www.e1evate.me								
N/A	N/A								
									TRUE
https://d2iq.com/	d2iq.com								
https://paintedrockcapitalgroup	paintedrockcapitalgroup	TRUE							
http://www.fabtech.net	www.fabtech.net	TRUE							
https://www.seodogs.com/	www.seodogs.com								
https://www.s-5.com	www.s-5.com	TRUE							
N/A									
http://www.palmetto.com	www.palmetto.com	TRUE							
http://www.talentjobs.ir	www.talentjobs.ir								
http://www.xcelenergy.com	www.xcelenergy.com	TRUE				TRUE			TRUE
http://www.k2-systems.com	www.k2-systems.com	TRUE							
http://bloomberg.com/compan	bloomberg.com								TRUE
https://e3lithium.ca/	e3lithium.ca				TRUE				
http://www.nabcep.org	www.nabcep.org	TRUE							TRUE
http://www.pam-membranas.c	pam-membranas.com.br								
http://www.iag.com.au	iag.com.au								
N/A									
http://www.daveyresourcegrou	www.daveyresourcegroup.com								
https://www.ibm.com/us-en	https://www.ibm.com								TRUE
https://www.s-5.com	www.s-5.com	TRUE							
http://www.ochsner.org/career	www.ochsner.org								

Company Classification Agent - AI Agent Structure

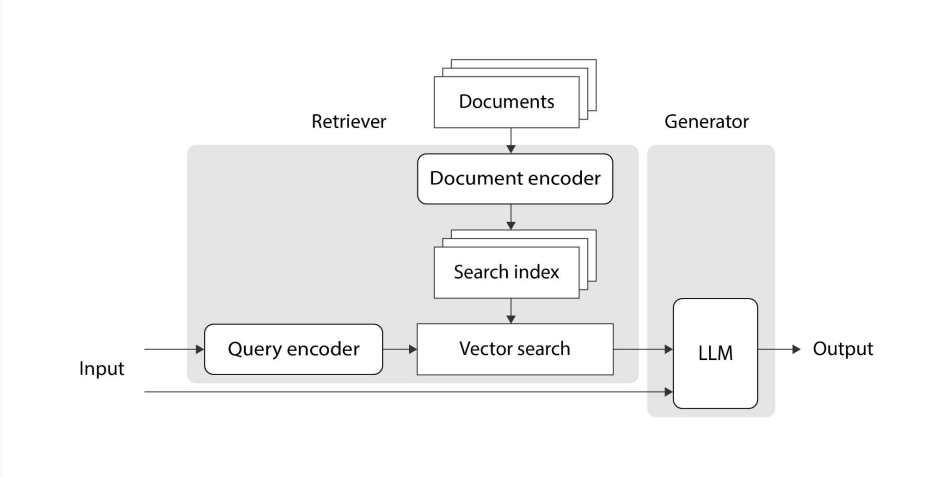


Crew AI framework:

- Accessible

RAG-based web scraping tool:

- Save token, reduce cost



Agents

Researcher

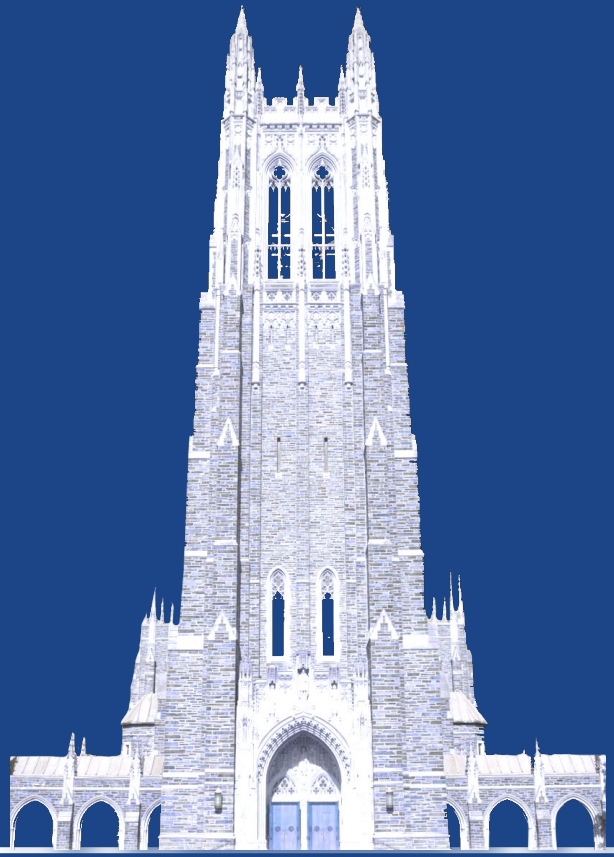
- Scrap original website
- Search for relevant website (About Us, Mission, What We Do)
- Scrap relevant website

Classifier

- Classify company into the categories provided
- Return a list

```
Relevant classifications for Painted Rock Capital Group are:
- **Energy Technology**: Solar, Energy Storage
- **Solar Project Category**: Commercial & Industrial, Utility-Scale
- **Solar Industry Role**: Financing, Power Producers, Operations and Maintenance, Developers
```

DIIG

Thank You

- Questions?
- Next Steps