# Strategic Decoupling:

Al-based analysis on Language-Action Nexus









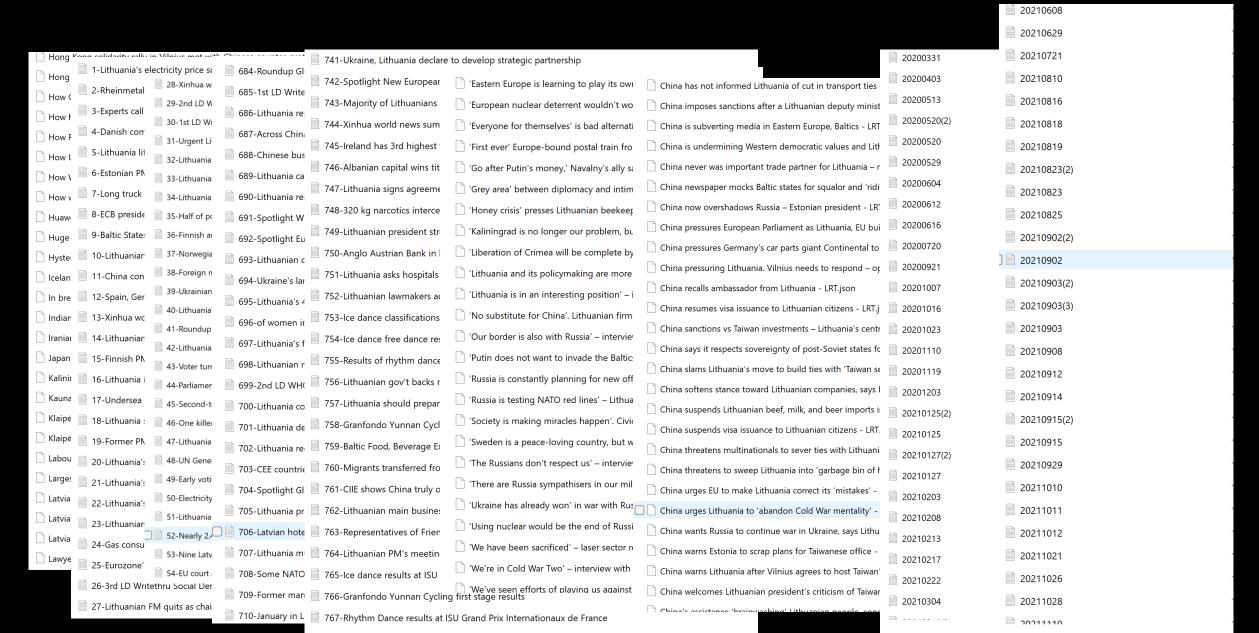
Annual Conference Best Conference Paper Finalist National AI Research Grant Sheen S. Levine
Otto Moensted Professor
DTU Copenhagen,
University of Texas, Dallas
& Columbia University

Mike M. Peng University of Texas, Dallas

Al and Big Data Masterclass In International Management
AoM 2025



## Statements & News



# What Are We Seeking?

- 1. What is the sentiment of China towards Lithuania?
- 2. What is the sentiment of Lithuania towards China?
- 3. What are topics discussed?
- 4. How do they change overtime?

#### For each statement or News:

Review context

Identify subjects

Identify sentiment and topic towards subjects

Consistency

# Why aspect-based sentiment analysis with LLMs?

e.g.,

Mike is sad, John is sad, you are happy because Mike is sad.

Three negative words (sad, sad, sad), one positive word (happy) By count words and its ratio, overall is negative sentiment.

But.. What if I care about your emotion?

# Why topic modelling with LLMs?

- News A: "US imposes new sanctions on Chinese tech firms over national security concerns, restricting chip exports." (Geopolitical Tech Conflict, US-> China)
- News B: "EU regulators approve sanctions against Russian oil imports, citing violations of international law." (Economic sanction on energy, EU-> Russia)

# Text-Embedding-3-large via Application Programming Interface

Sampling

Prompt tuning

Aspect-based sentiment analysis

Dynamic topic extraction

Validation

target subjects= ["Lithuania", "Taiwan", "US", "China"]

## Aspect-based analysis

#### Sentiment &

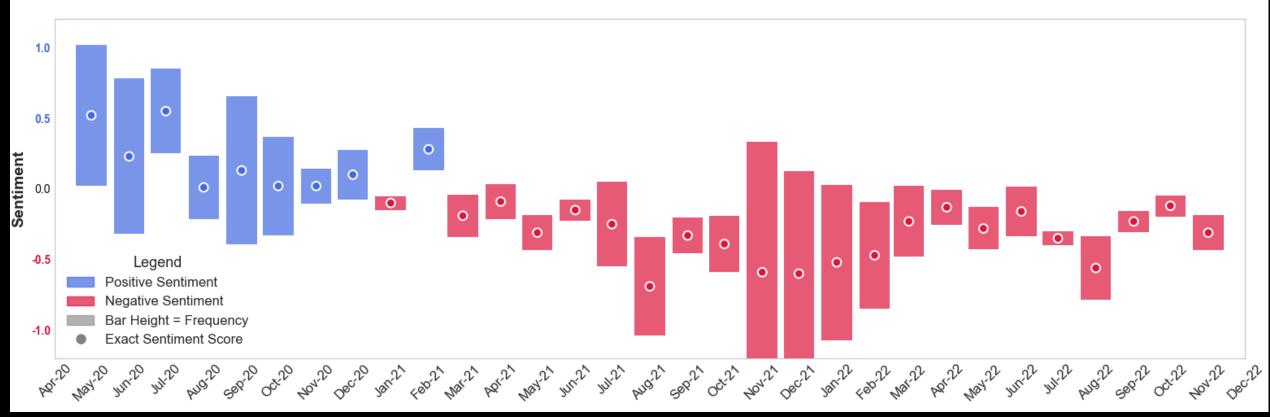
```
if target_subject:
    analysis_target = f"specifically towards
'{target_subject}'"
else:
    analysis_target = "for the overall text"
```

# Prompt-guided topic extraction

Your task is to determine if the following topic phrase is specifically about '{target\_subject}' or if it's a more general concept.

# Findings





# How to conduct "topic modelling" with LLMs?

- News A: GPT Summary -> Topics + Reasons + Quote
- News B: GPT Summary -> Topics + Reasons + Quote
- News C: GPT Summary -> Topics + Reasons + Quote
- News D: GPT Summary -> Topics + Reasons + Quote
- •
- News N: GPT Summary -> Topics + Reasons + Quote

### Vectorization: Text → Value

#### Chunks

	Women	Men	Royal	Poor	•••	
Queen	1	0	1	0	•••	(1,0,1,0)
King	0	1	1	0	•••	(0, 1,1,0)

Vectors

### Topic Distance

- Distance(Embedding A, Embedding B) -> Similarity Score (A,B)
- Distance(Embedding B, Embedding C) -> Similarity Score (B,C)
- Distance(Embedding C, Embedding A) -> Similarity Score (C,A)



#### Validation:

Randomly pick to compare with human experts

Compare with LIWC, full file sentiment analysis

Further compare with Gemini models & Retrieval augmented generation

# Alternative Method



Human coding

Bag of words

Al-based language analysis

# Human coding







# Bag-of-Words





Fit: Is AI the right tool for the job?

- Identify advantage
- Match capacity to task

**Use**: The standard is transparency

Document & disclose

Validate: Trust, but verify

- Start with test
- Benchmark performance



### Takeaways









#### How about LLMs?

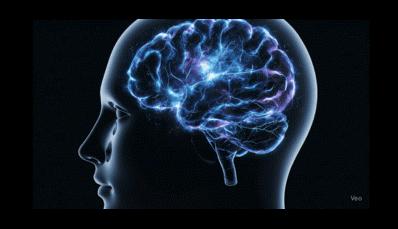
#### Advantages (5C)

- Context
- Consistent
- Comprehensive
- Cheap
- Composable

#### Concerns

- Black box
- Hallucination
- Update quick

...?



```
def get_sentiment_with_llm(text_content, target_subject=None):
    if target_subject:
        analysis_target = f"specifically towards '{target_subject}'"
    else:
        analysis_target = "for the overall text"
        system_prompt = f"""You are a sentiment analysis engine. Your task is to analyze the sentiment {analysis_target}.
You MUST respond with a valid JSON object with two keys:
1. "score": A numerical score from -1.0 (very negative) to 1.0 (very positive), with 0 being neutral.
2. "reasoning": A brief explanation (1-2 sentences) for your score, citing specific phrases or ideas from the text."""
```

```
try:
    response = client.chat.completions.create(
        model=AZURE_DEPLOYMENT_NAME,
        messages=[
            {"role": "system", "content": system_prompt},
            {"role": "user", "content": text_content}
        ],
        response_format={"type": "json_object"},
        # --- USER-REQUESTED PARAMETERS ---
        temperature=0.7,
        top_p=1.0, # top_p=1 is the default, but we set it
        n=num_completions
```

```
# --- 1. SETUP (No changes needed) ---
CHAT_API_KEY = os.getenv("AZURE_CHAT_API_KEY", "1eVHSu0ETLN1V2kDngWXMZDp9CxXi8GxvT9DYRuWKVMCZH
CHAT_API_ENDPOINT = os.getenv("AZURE_CHAT_API_ENDPOINT", "https://big-agi.openai.azure.com/")
CHAT_DEPLOYMENT_NAME = os.getenv("AZURE_CHAT_DEPLOYMENT_NAME", "gpt-4.1") # e.g., "MyCompany-G

# Get these values from the Azure resource where your EMBEDDING model is deployed.

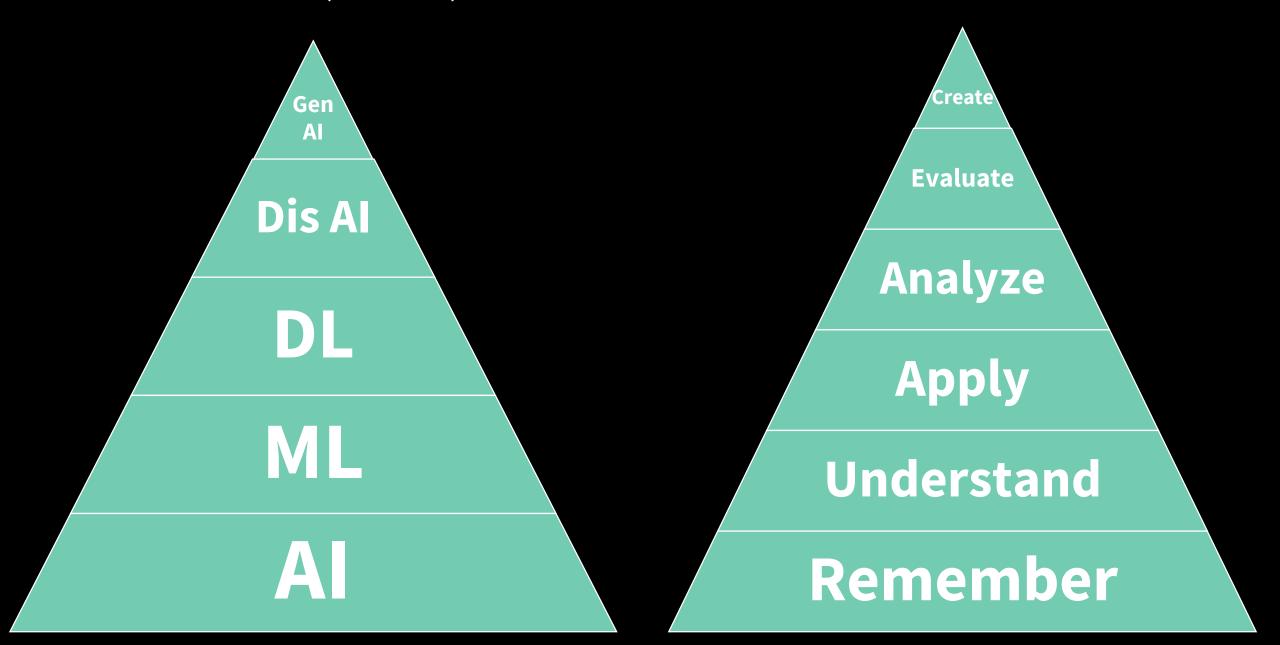
EMBEDDING_API_KEY = os.getenv("AZURE_EMBEDDING_API_KEY","1eVHSu0ETLN1V2kDngWXMZDp9CxXi8GxvT9DY
EMBEDDING_API_ENDPOINT = os.getenv("AZURE_EMBEDDING_API_ENDPOINT", "https://big-agi.openai.azu
EMBEDDING_DEPLOYMENT_NAME = os.getenv("AZURE_EMBEDDING_DEPLOYMENT_NAME", "text-embedding-3-lar
```

```
def classify_topic_with_llm(topic_string, aspect_name, client): 1usage
   system_prompt = f"""You are a topic classification expert. Your task is to determine if the followin
For example, if the aspect is 'Lithuania':
- "Lithuanian-Polish relations" is SPECIFIC.
- "NATO defense strategy" is SPECIFIC (as Lithuania is a key part of it).
- "Economic inflation" is GENERAL.
- "Human rights" is GENERAL.
Your response MUST be a valid JSON object with a single key "is_specific_to_aspect", which must be a boo
   try:
      response = client.chat.completions.create(model=CHAT_DEPLOYMENT_NAME,
                                            messages=[{"role": "system", "content": system_prompt}
                                                    {"role": "user",
                                                     "content": f"Topic phrase: \"{topic_string}
                                            response_format={"type": "json_object"}, temperature=1
      return json.loads(response.choices[0].message.content).get("is_specific_to_aspect", False)
 # Step 4: Get embeddings and run t-SNE on ALL themes from ALL periods
  print("\n\nGetting embeddings for all generated themes...")
 for theme in all_themes_data:
       theme['embedding'] = get_embedding(theme['theme_description'])
  all_themes_data = [t for t in all_themes_data if t.get('embedding')]
```

# What I wish I know before?

- Documentation: Always record your API version, temperature setting, and threshold score, and definition.
- Always experiment with small samples, to see how does code work, and randomly verify outcome.
- Test for a "best setting" and then run with full dataset.
- Clear and precise prompt.
- Validation: compare with alternative tools and human experts.

# What is AI, ML, DL and Gen AI?



#### Al in Management Study

- Measuring the "Unmeasurable"
- ✓ measure team cognitive diversity (Lix et al, 2022)
- ✓ measure organizational identification (Yang et al., 2025)

- Identifying Novelty and Predicting the Future
- ✓ To find 'prescient' ideas (Vicinanza et al., 2023)

- Uncovering Hidden Social Dynamics Through Linguistic Style
- ✓ measure post-merger integration (Bhatt et al., 2022)
- ✓ create dynamic measures from static surveys (Lu et al., 2024)

# Word Embeddings

- Represents words as vectors, so we can calculate the 'distance' or 'similarity' in meaning between words, sentences, or people.
- Measuring discursive diversity in teams by calculating the average semantic distance between members' Slack messages (Lix et al., 2022).
- Measuring organizational identification by tracking the similarity between a person's concept of "I" and "we" in their emails (Yang et al., 2025).

#### Deep Learning Models (Measuring Novelty)

- What it does: Understands context so well it can tell you how 'surprising' or 'predictable' a given sentence is. This is called 'perplexity'.
- Identifying prescient ideas by finding text that was highly surprising at the time it was written but became predictable years later (Vicinanza et al., 2023).