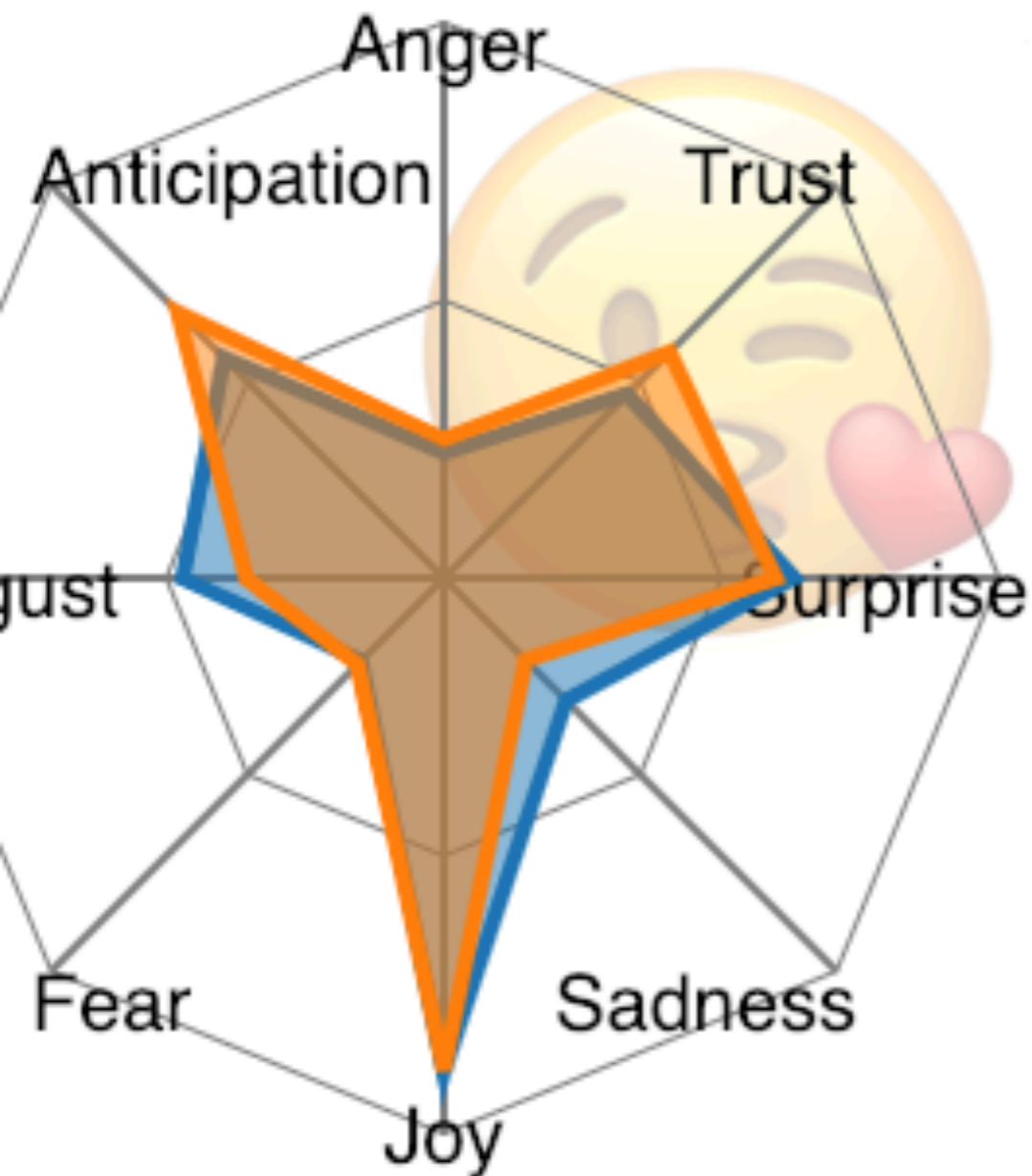


# Emoji-Emotion Association: (Dis-)Similarities Across Twitter and Weibo



Mingyang Li  
Sharath Chandra Guntuku  
Vinit Jakhetiya  
Lyle H. Ungar

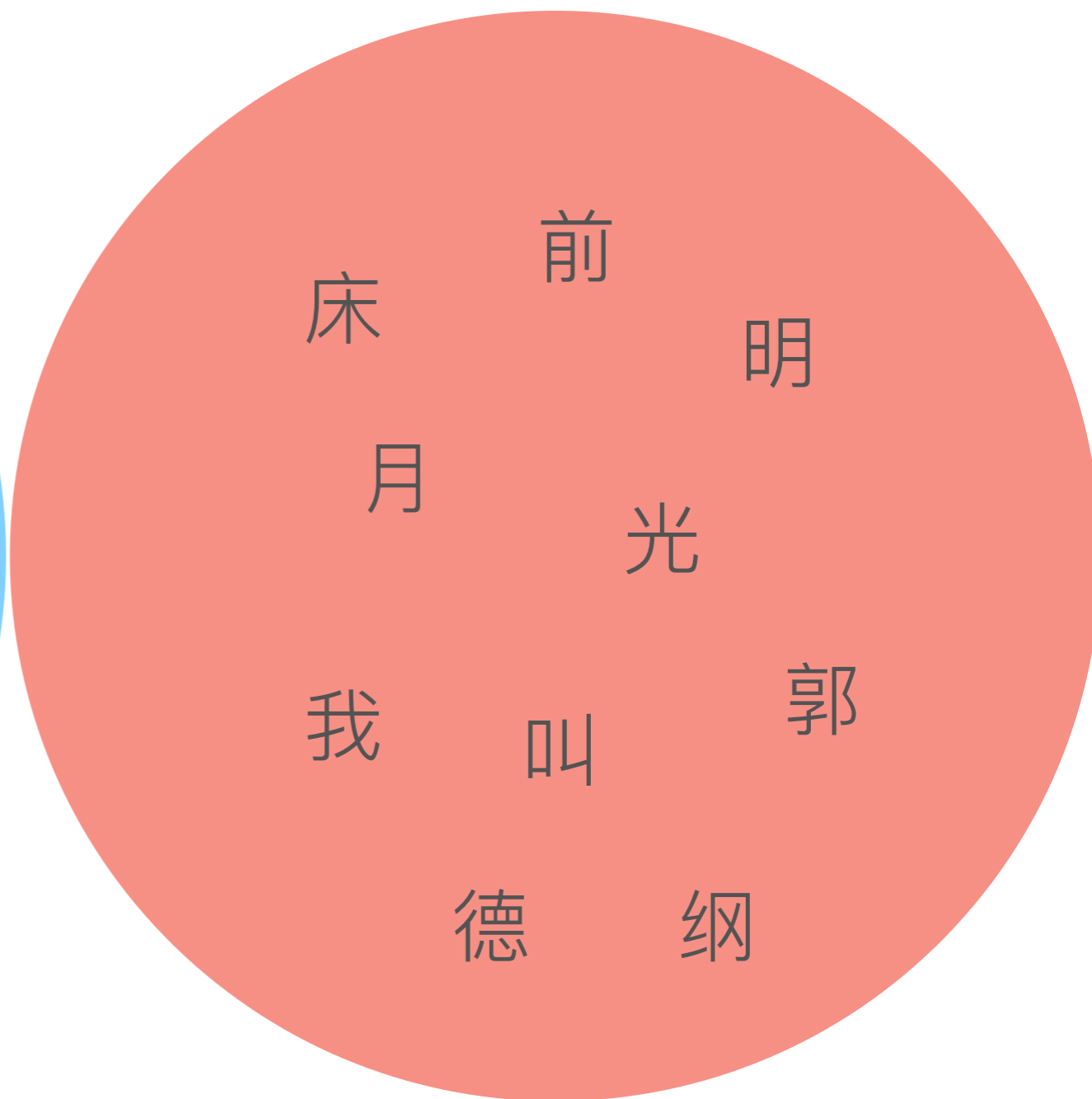
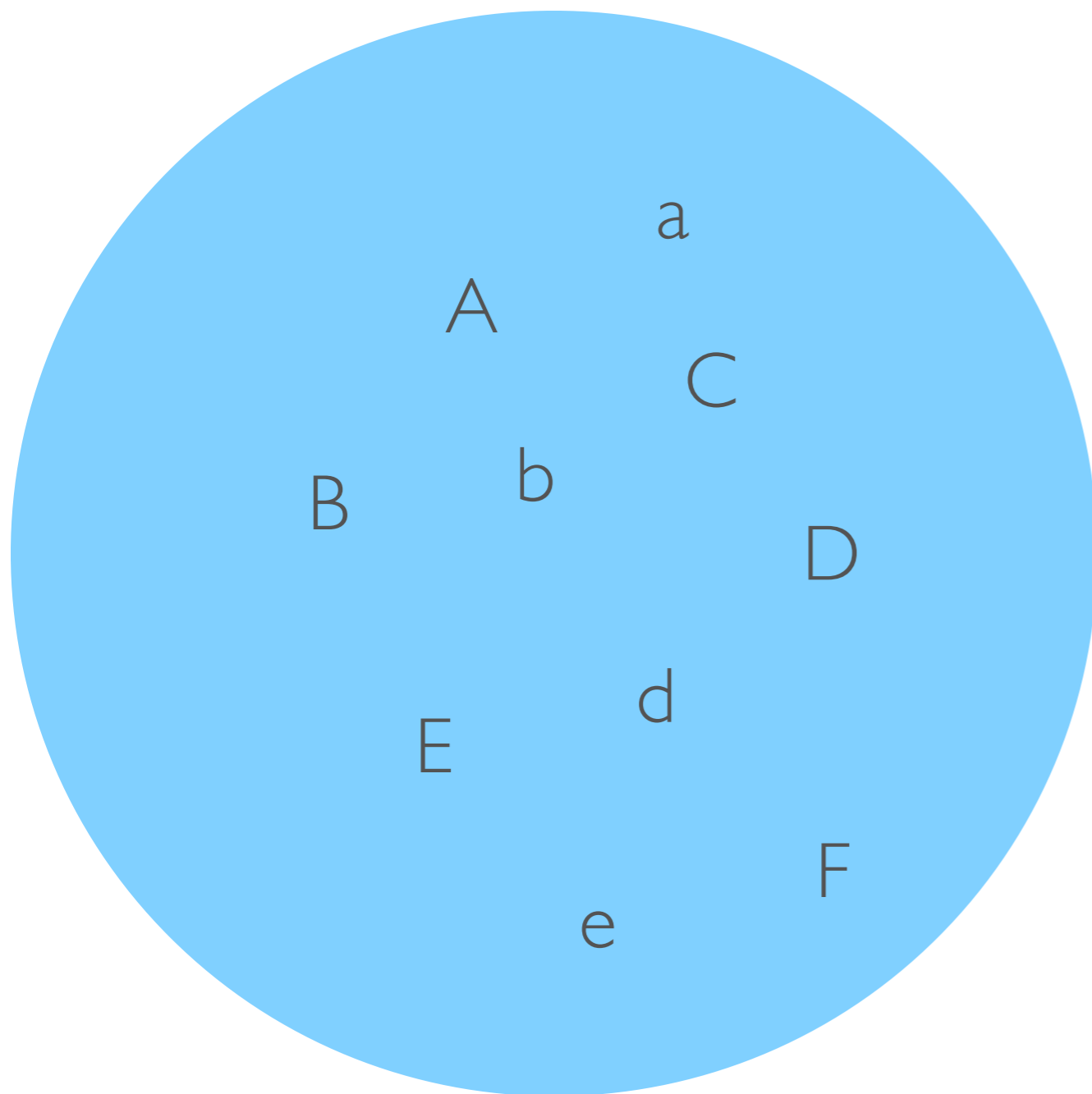
# Table of Contents

1. **Two main ingredients:** Emojis and EmoLex
2. **Method:** Corpora, word2vec, average by group, and compare
3. **Results:** Anatomy and panorama
4. **Discussion:** Case study and correlation/similarity
5. **What's Next:** More topics and more emoticons

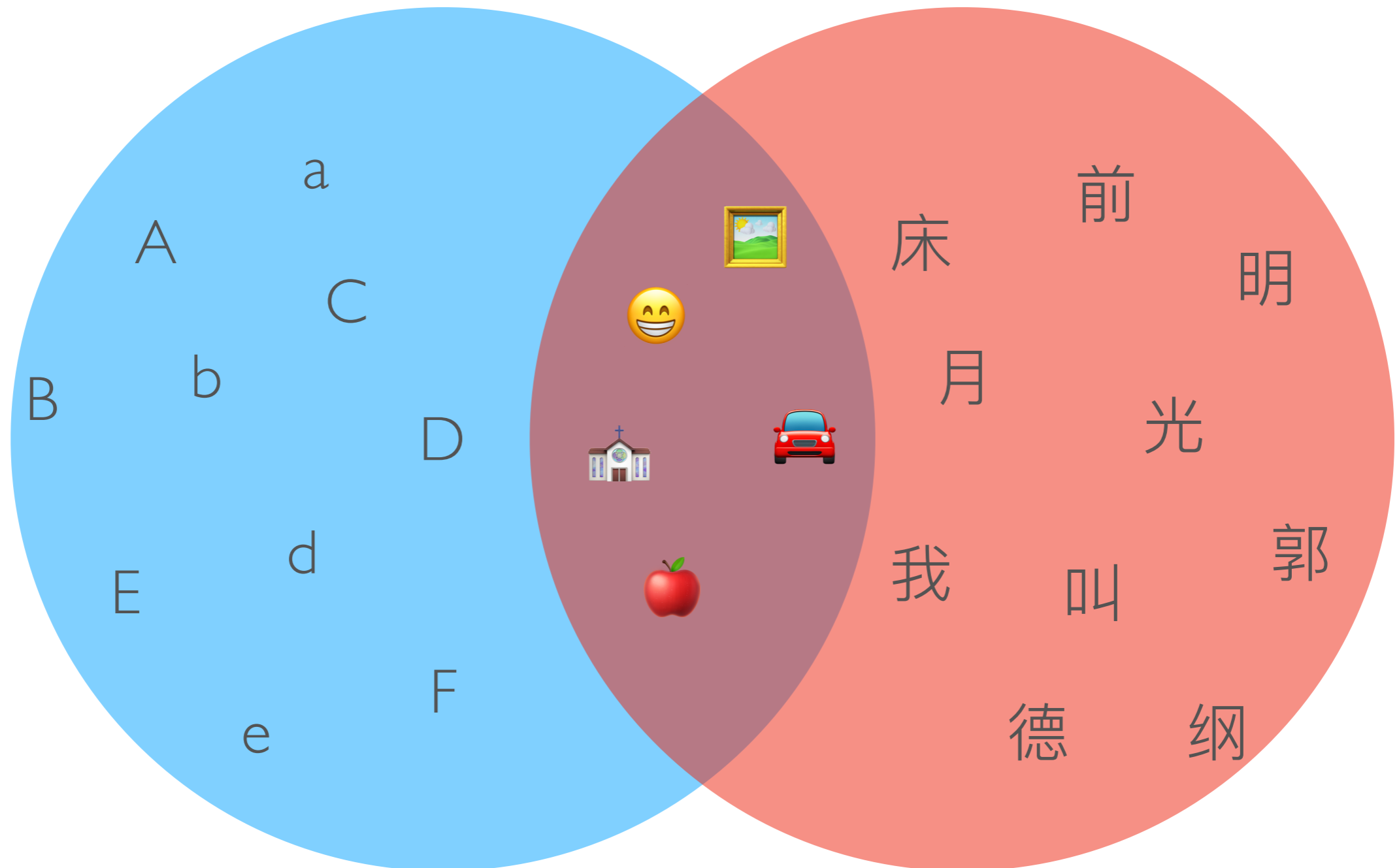
# Table of Contents

1. **Two main ingredients:** Emojis and EmoLex
2. **Method:** Corpora, word2vec, average by group, and compare
3. **Results:** Anatomy and panorama
4. **Discussion:** Case study and correlation/similarity
5. **What's Next:** More topics and more emoticons

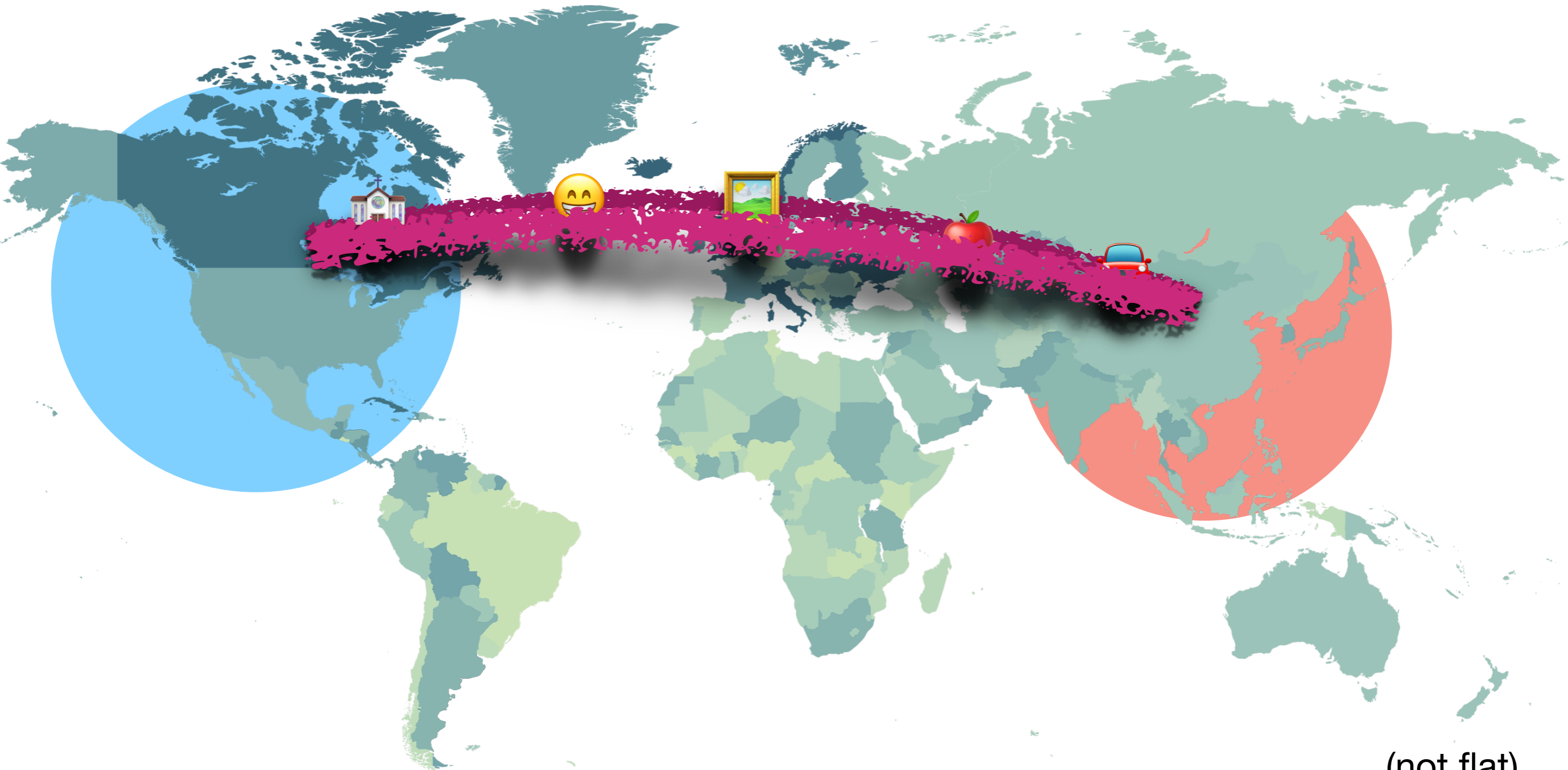
# English and Chinese share no character...



# ... until Emojis were invented.



# That was our first bridge between East and West.

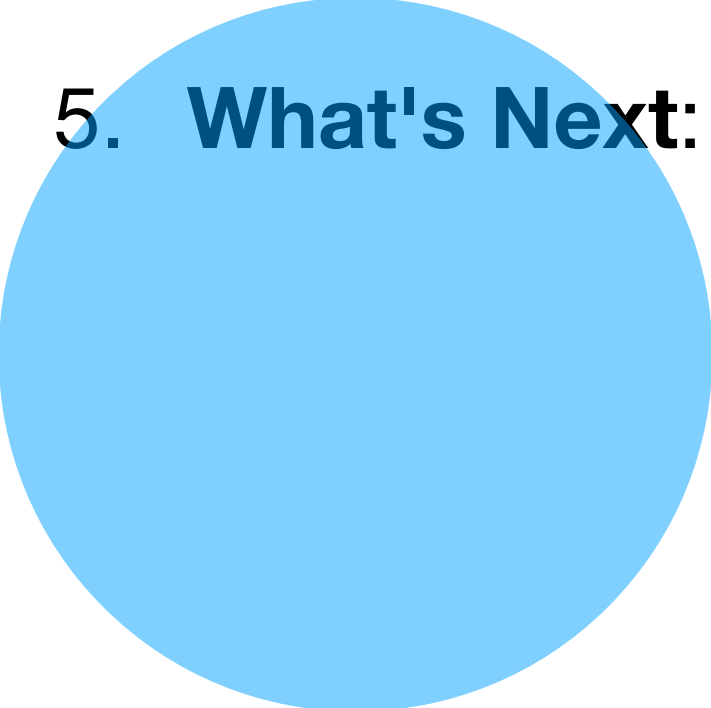


(not flat)



# Table of Contents



1. **Two main ingredients:** Emojis and EmoLex
  2. **Method:** Corpora, word2vec, average by group, and compare
  3. **Results:** Anatomy and panorama
  4. **Discussion:** Case study and correlation/similarity
  5. **What's Next:** More topics and more emoticons
- 



# How did we do it?

1. 🇺🇸

1. Train *word2vec* on **U.S. Twitter** corpus.
2. Group **English** words by emotion; take average. → *Emotion Vectors*
3. Calculate vectorial similarity between each pair of **Emoji** and *Emotion*. → *Emoji-Emotion Similarity Table in English*

2. 🇨🇳

1. Train *word2vec* on **Chinese Weibo** corpus.
2. Group **Chinese** words by emotion; take average. → *Emotion Vectors*
3. Calculate vectorial similarity between each pair of **Emoji** and *Emotion*. → *Emoji-Emotion Similarity Table in Chinese*

3. Compare *Table in English* with *Table in Chinese*.

# 1 Collect Corpora

microblog posts from US and China, separately



twitter



# 2 word2vec: Vector View

maps tokens to vectors



"抽烟" = [6, 3, 6, 7, 8, 4, 2, 0, 8, 0]

"!" = [2, 7, 0, 5, 8, 9, 3, 4, 6, 2]

"喝酒" = [6, 4, 8, 0, 2, 3, 8, 9, 7, 1]

"🚗" = [4, 6, 0, 9, 1, 2, 4, 5, 1, 8]

"烫头" = [7, 7, 4, 1, 2, 6, 5, 3, 3, 4]

"😊" = [5, 2, 4, 6, 7, 9, 4, 7, 9, 3]

...

"monkey" = [1, 2, 3, 7, 3, 5, 8, 0, 5, 2]

"😊" = [5, 2, 1, 2, 8, 8, 5, 7, 3, 3]

"Emoji" = [2, 5, 6, 7, 1, 5, 4, 3, 8, 7]

"great" = [7, 8, 9, 9, 2, 2, 5, 7, 6, 1]

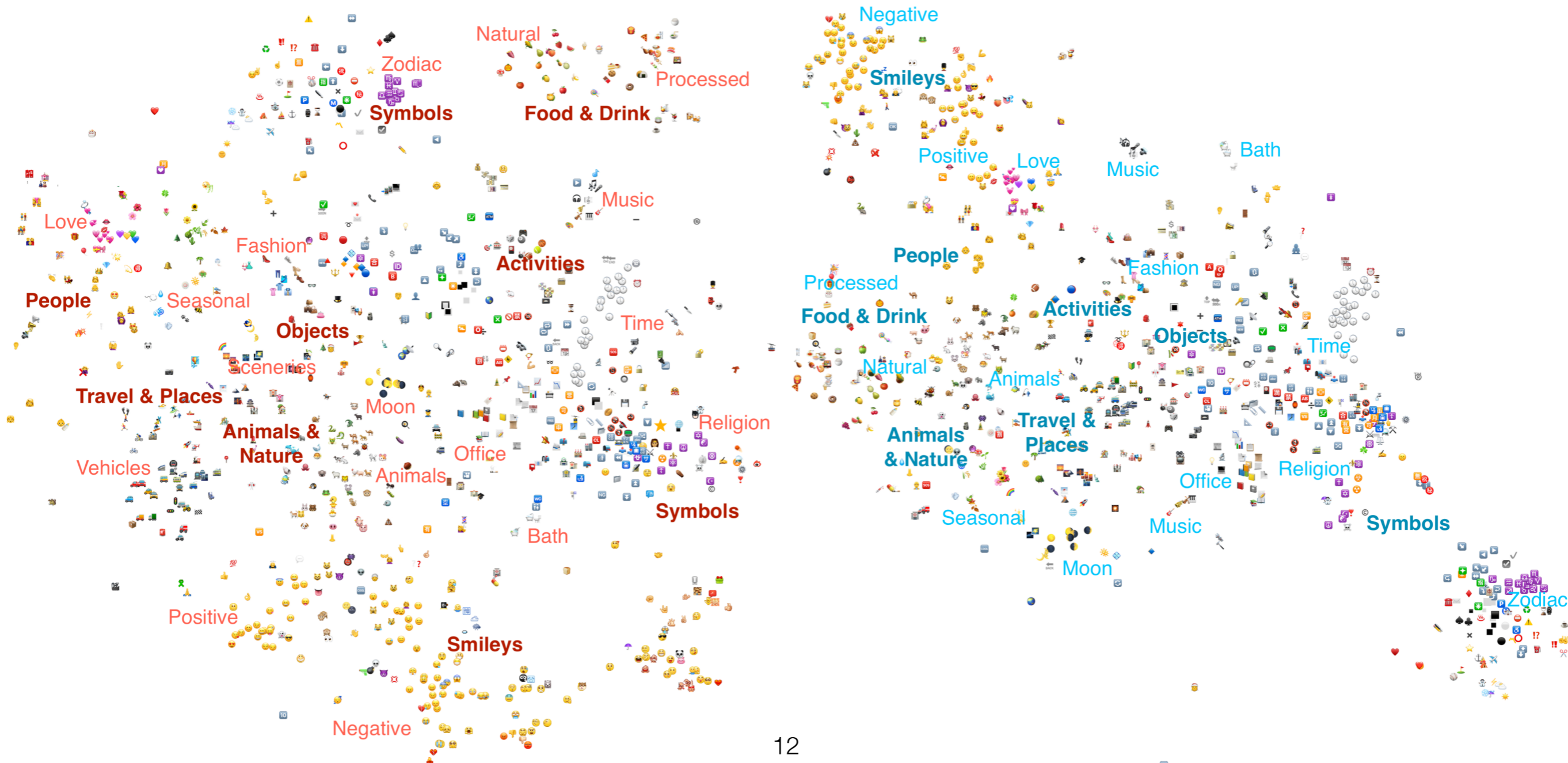
"!" = [9, 3, 4, 6, 2, 1, 7, 0, 0, 8]

"🚗" = [3, 3, 7, 5, 9, 8, 2, 1, 1, 2]

...

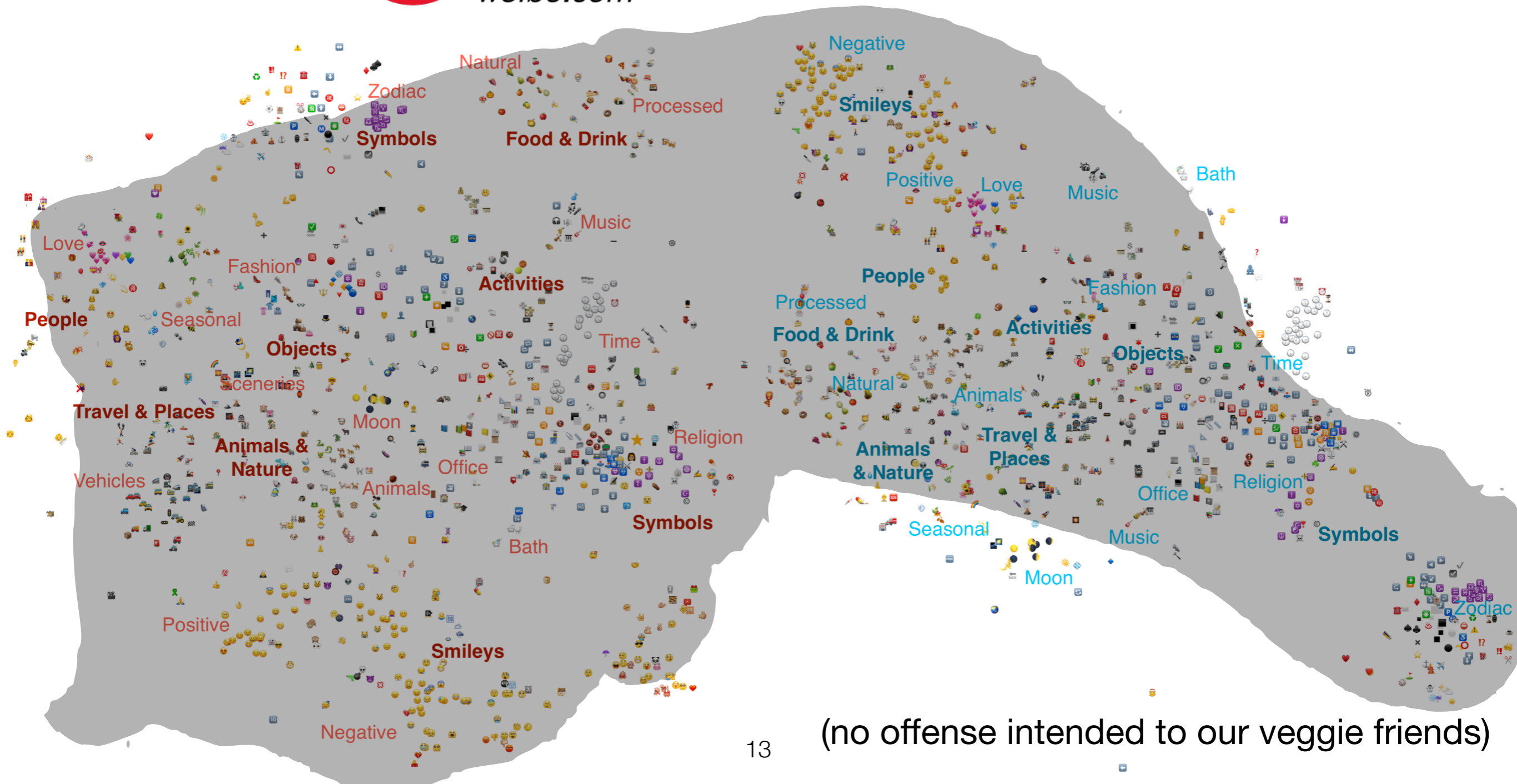
# 2 word2vec: t-SNE View

(almost a convention)



# 2 word2vec: t-SNE View

similarity to a chicken thigh was unintentional



# 2 word2vec: Vector View

maps tokens to vectors



"抽烟" = [6, 3, 6, 7, 8, 4, 2, 0, 8, 0]

"!" = [2, 7, 0, 5, 8, 9, 3, 4, 6, 2]

"喝酒" = [6, 4, 8, 0, 2, 3, 8, 9, 7, 1]

"🚗" = [4, 6, 0, 9, 1, 2, 4, 5, 1, 8]

"烫头" = [7, 7, 4, 1, 2, 6, 5, 3, 3, 4]

"😊" = [5, 2, 4, 6, 7, 9, 4, 7, 9, 3]

...

"monkey" = [1, 2, 3, 7, 3, 5, 8, 0, 5, 2]

"😊" = [5, 2, 1, 2, 8, 8, 5, 7, 3, 3]

"Emoji" = [2, 5, 6, 7, 1, 5, 4, 3, 8, 7]

"great" = [7, 8, 9, 9, 2, 2, 5, 7, 6, 1]

"!" = [9, 3, 4, 6, 2, 1, 7, 0, 0, 8]

"🚗" = [3, 3, 7, 5, 9, 8, 2, 1, 1, 2]

...

# 3 Group by Emotion

take average; call them "emotion vectors"



**Anger** = [4, 5, 7, 3, 2, 6, 7, 9, 8, 6]  
**Fear** = [0, 0, 8, 5, 7, 6, 8, 5, 2, 0]  
**Joy** = [1, 5, 3, 7, 6, 2, 9, 7, 4, 7]  
"😊" = [5, 2, 4, 6, 7, 9, 4, 7, 9, 3]

...



**Joy** = [9, 1, 4, 3, 2, 5, 4, 1, 3, 2]  
"😊" = [5, 2, 1, 2, 3, 8, 5, 7, 3, 3]  
**Anger** = [7, 8, 9, 5, 4, 1, 2, 8, 3, 4]  
**Fear** = [5, 4, 1, 2, 4, 4, 4, 7, 9, 9]


...




# 4 Compute Vector Similarities

between each pair of emotion and emoji on each corpus



|       |  |  |  | ... |
|-------|---|---|---|-----|
| Anger | -0.8  | 0.01  | 0.02  | ... |
| Fear  | -0.6  | 0.05  | -0.01   | ... |
| Joy   | 0.9   | -0.02   | 0.03  | ... |
| ...   | ...   | ...   | ...   | ... |



|       |  |  |  | ... |
|-------|---|---|---|-----|
| Anger | -0.7  | 0.02  | -0.02   | ... |
| Fear  | -0.7  | 0.04  | 0.03  | ... |
| Joy   | 0.8   | -0.01   | -0.01   | ... |
| ...   | ...   | ...   | ...   | ... |



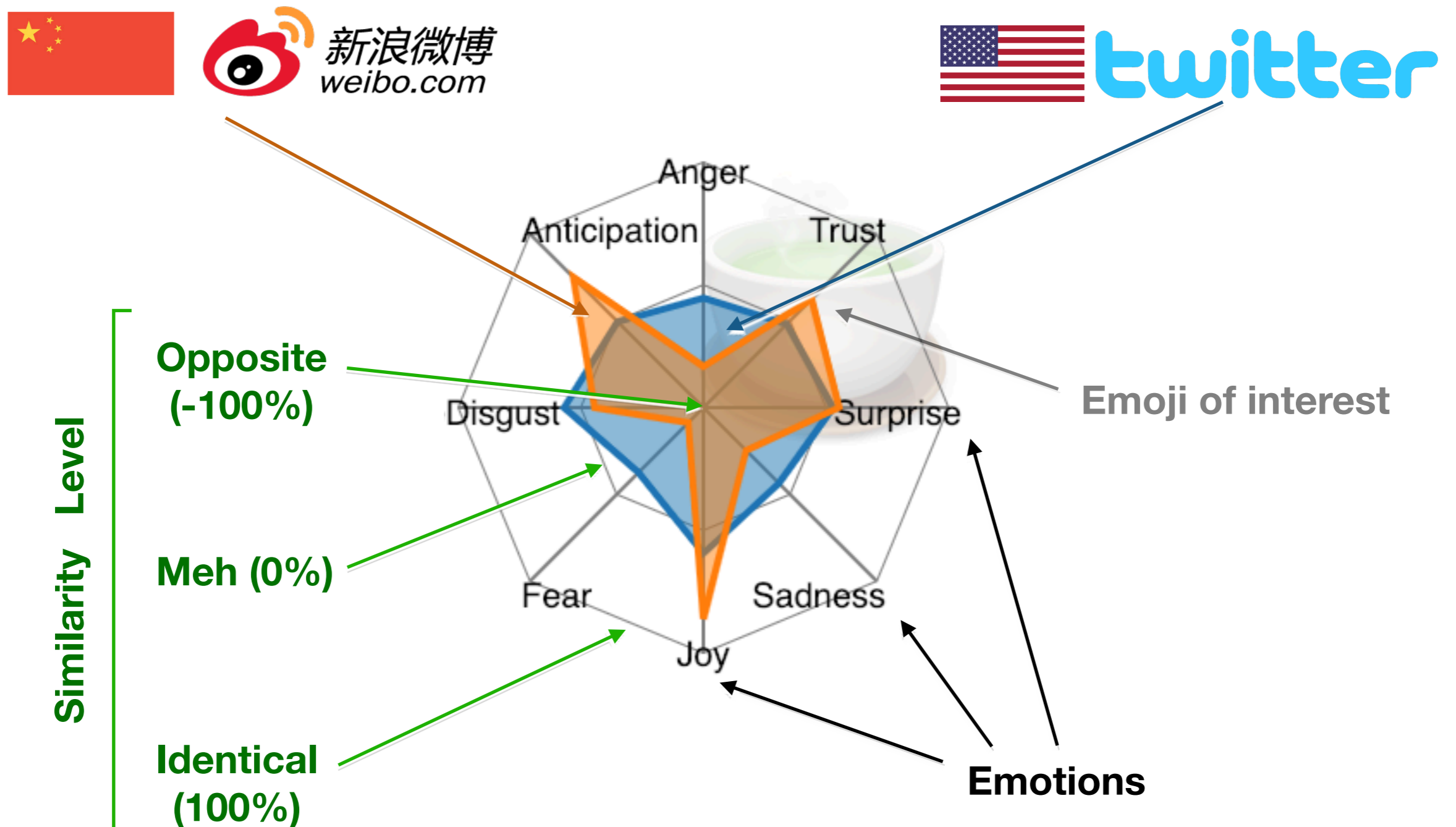
# Table of Contents

1. **Two main ingredients:** Emojis and EmoLex
2. **Method:** Corpora, word2vec, average by group, and compare
3. **Results:** Anatomy and panorama
4. **Discussion:** Case study and correlation/similarity
5. **What's Next:** More topics and more emoticons



# 4 Compare: Anatomy

"Tea correlates with positive emotions for Weibo users."



# 4 Compare: Full View top-5 most different

← → × 🔒 <https://unicode.org/emoji/charts/full-emoji-list.html>

**Unicode** **Emoji Charts**

## Full Emoji List, v12.0

[Index & Help](#) | [Images & Rights](#) | [Spec](#) | [Proposing Additions](#)

This chart provides a list of the Unicode emoji characters and sequences outlined images; their images may show as a group with “...” before and

Emoji with skin-tones are not listed here: see [Full Skin Tone List](#).

For counts of emoji, see [Emoji Counts](#).

While these charts use a particular version of the [Unicode Emoji data file](#)

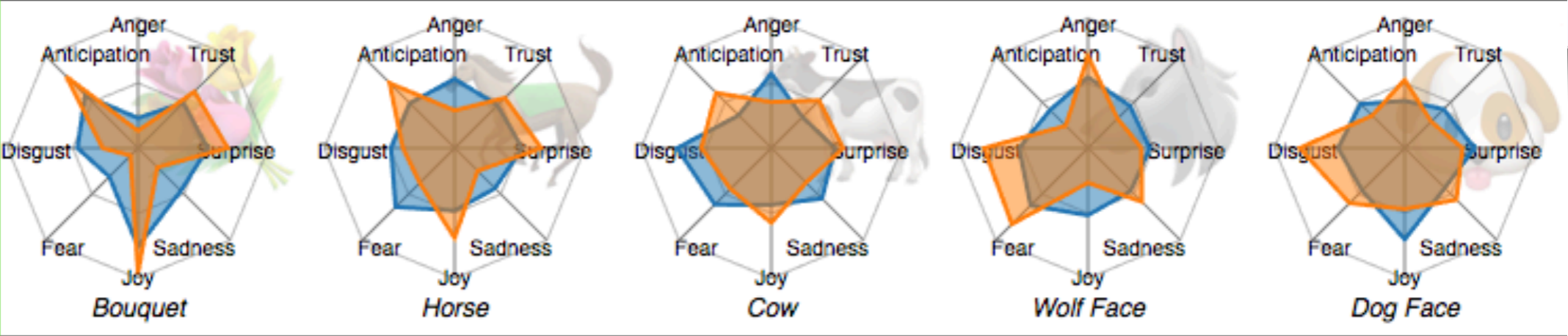
**Smileys & Emotion**

**face-smiling**

| No | Code                    | Browser | Appl | Goog | FB | Wind |
|----|-------------------------|---------|------|------|----|------|
| 1  | <a href="#">U+1F600</a> | 🤪       | 😊    | 😄    | 😁  | 😁    |

←

**Animals & Nature**



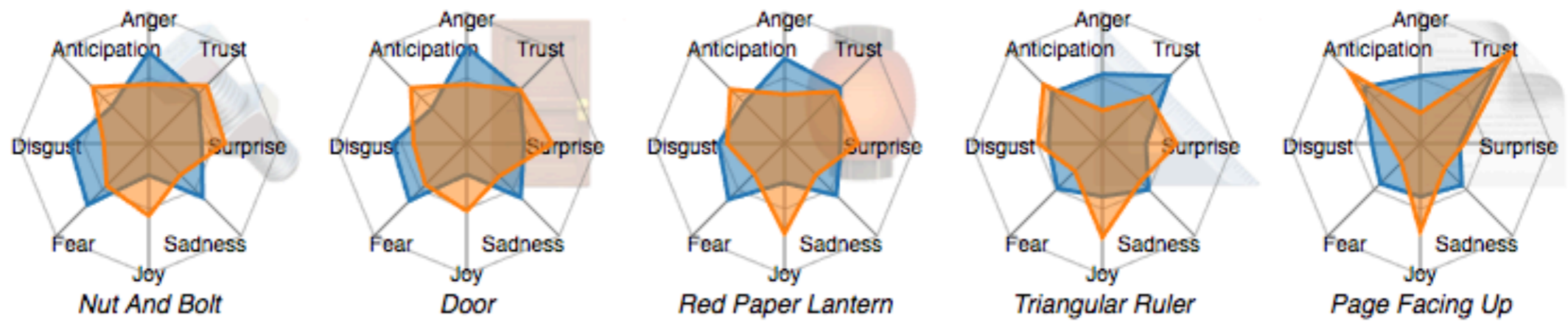
**Food & Drink**



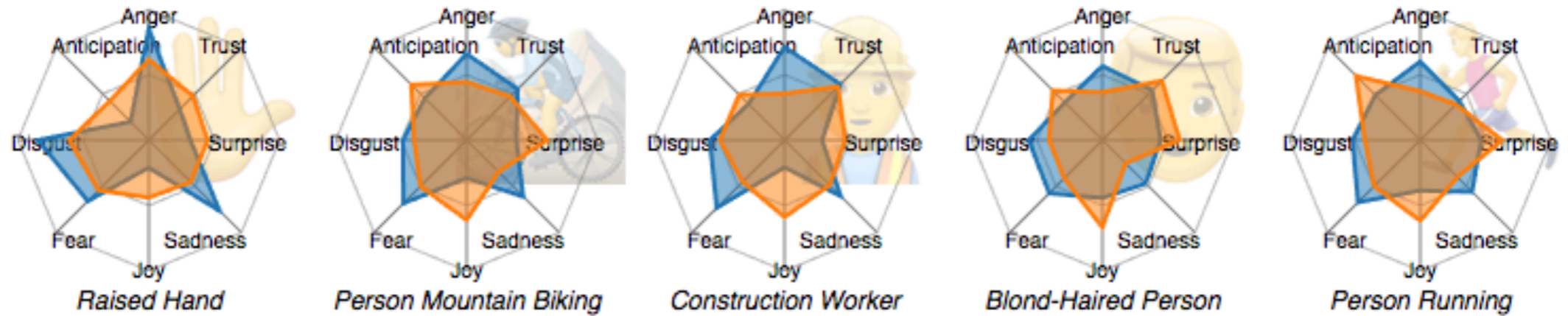
**Objects**



## Objects



## People & Body



## Smileys & Emotion



## Travel & Places



# Table of Contents

1. **Two main ingredients:** Emojis and EmoLex
2. **Method:** Corpora, word2vec, average by group, and compare
3. **Results:** Anatomy and panorama
4. **Discussion:** Case study and correlation/similarity
5. **What's Next:** More topics and more emoticons

# Case Study of 🍷 🍷👁️ 🍷👁️👁️ 🍷👁️👁️👁️

"Kissing Face"

"Kissing Face With Smiling Eyes"

"Kissing Face With Closed Eyes"

"Face Blowing a Kiss"



- 3-shaped lips



- Curly eyes
- 3-shaped lips



- Blush
- Curly eyes
- 3-shaped lips



- Wink and ❤️
- Blush
- 3-shaped lips
- Curly eyes



# Back to the Good Part

The pattern  or  reflects correlation between emotions.

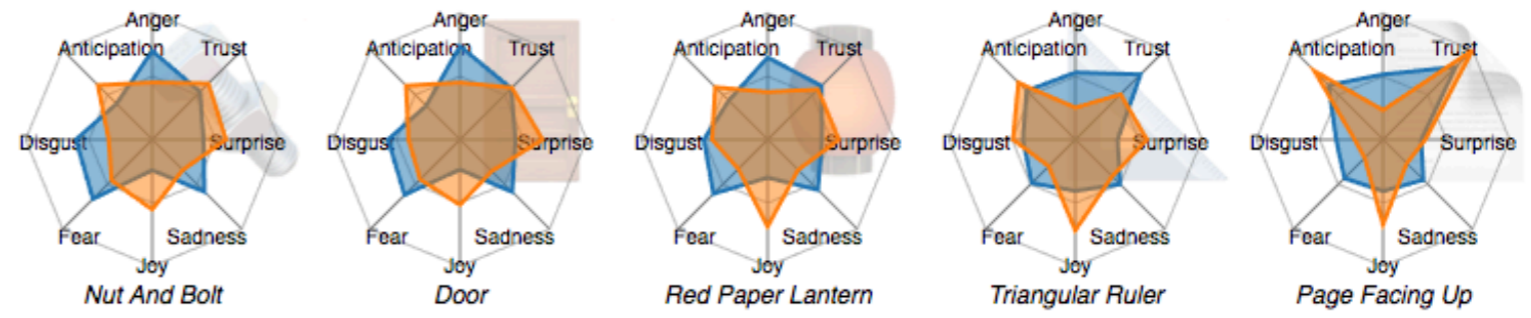
## Animals & Nature



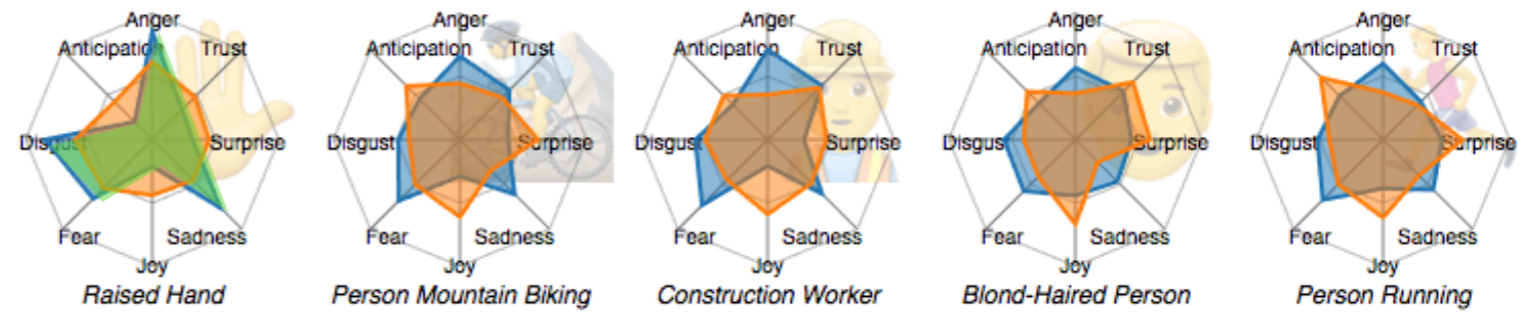
## Food & Drink



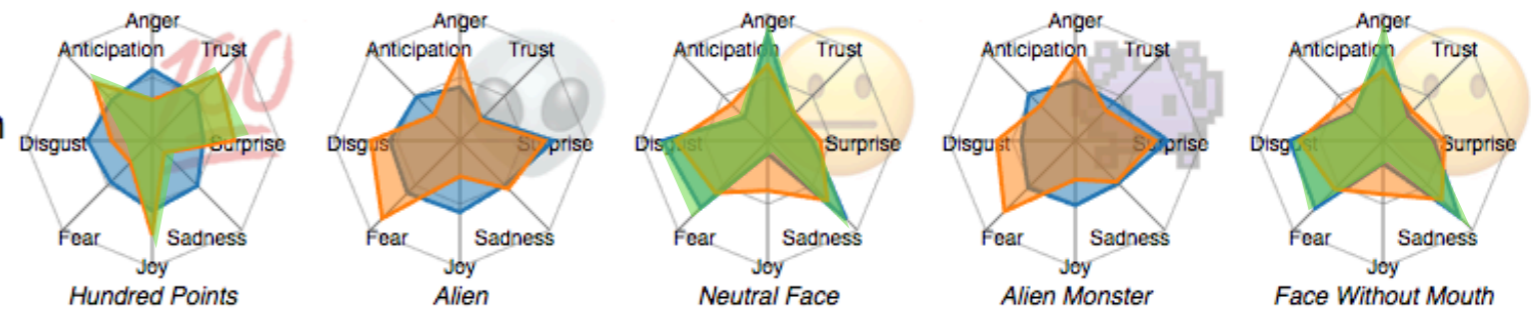
## Objects



## People & Body



## Smileys & Emotion



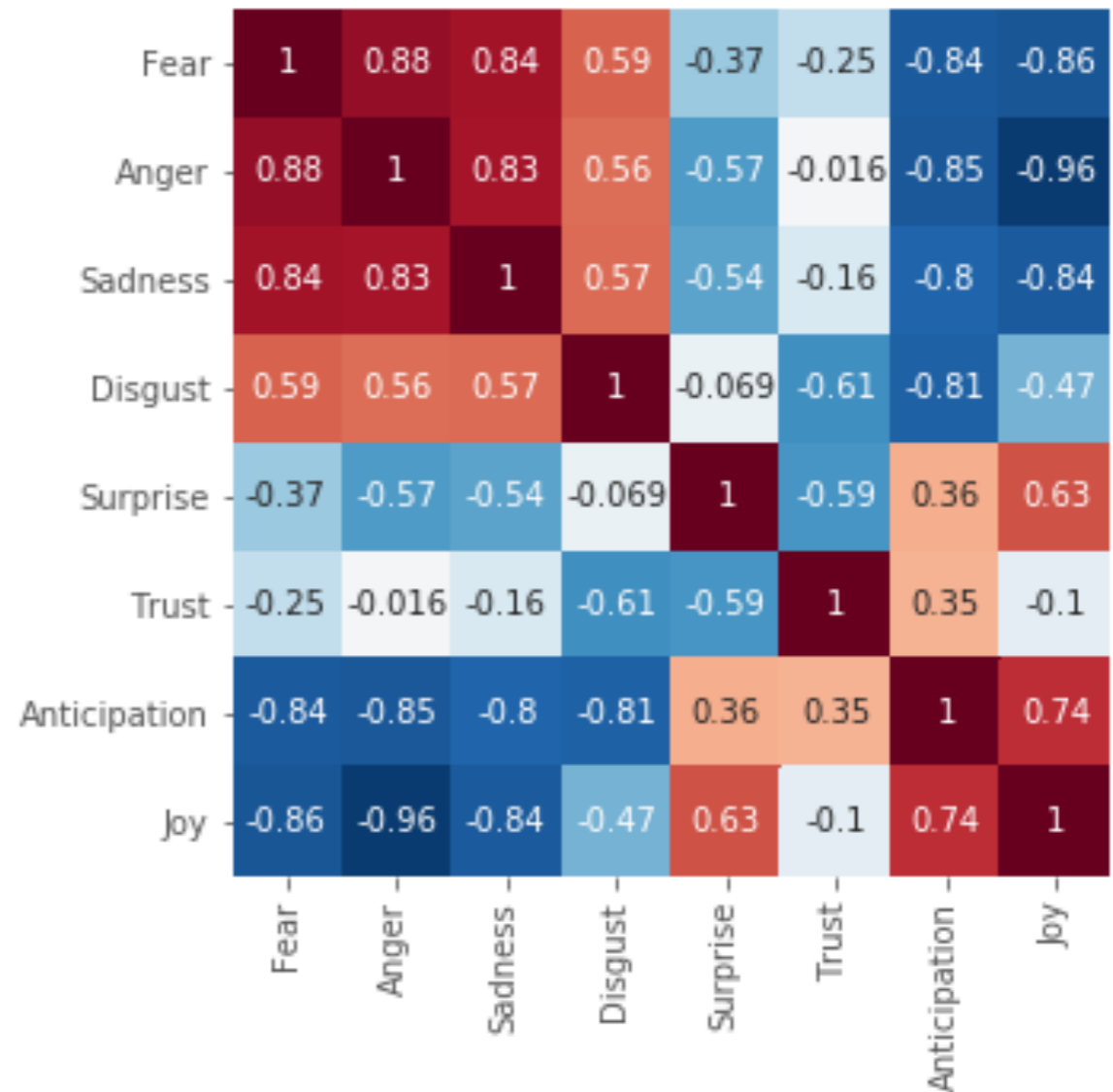
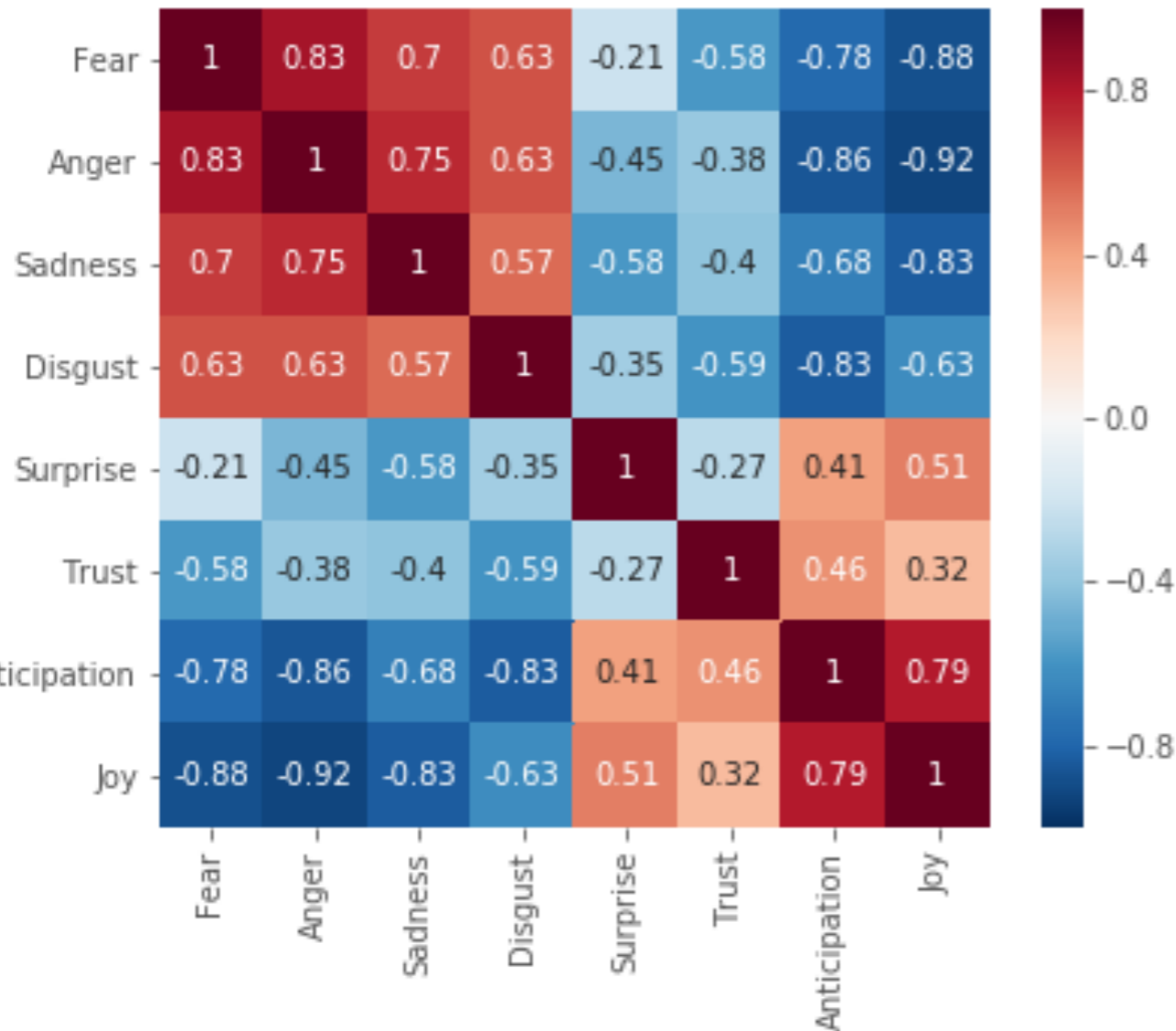
## Travel & Places





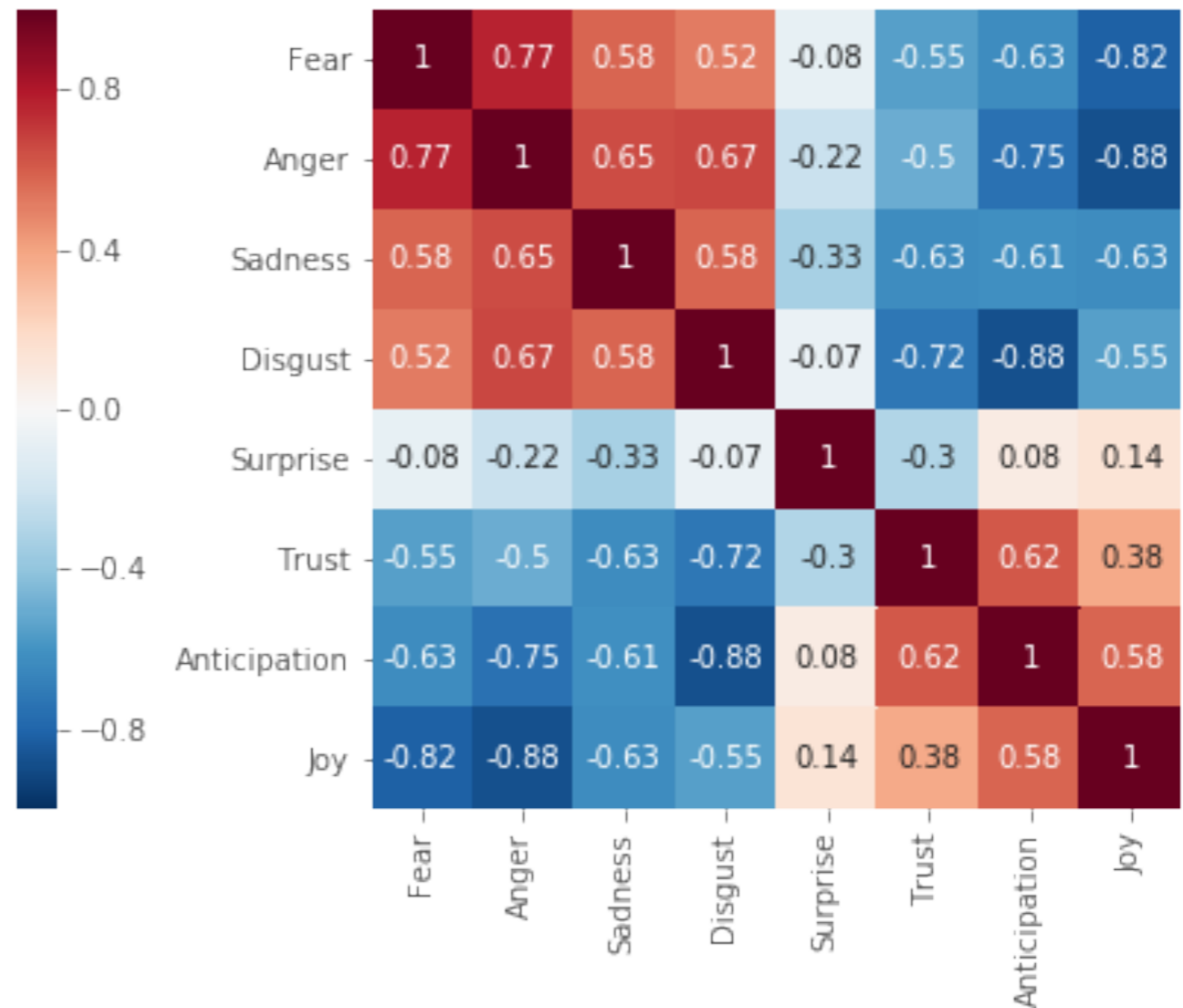
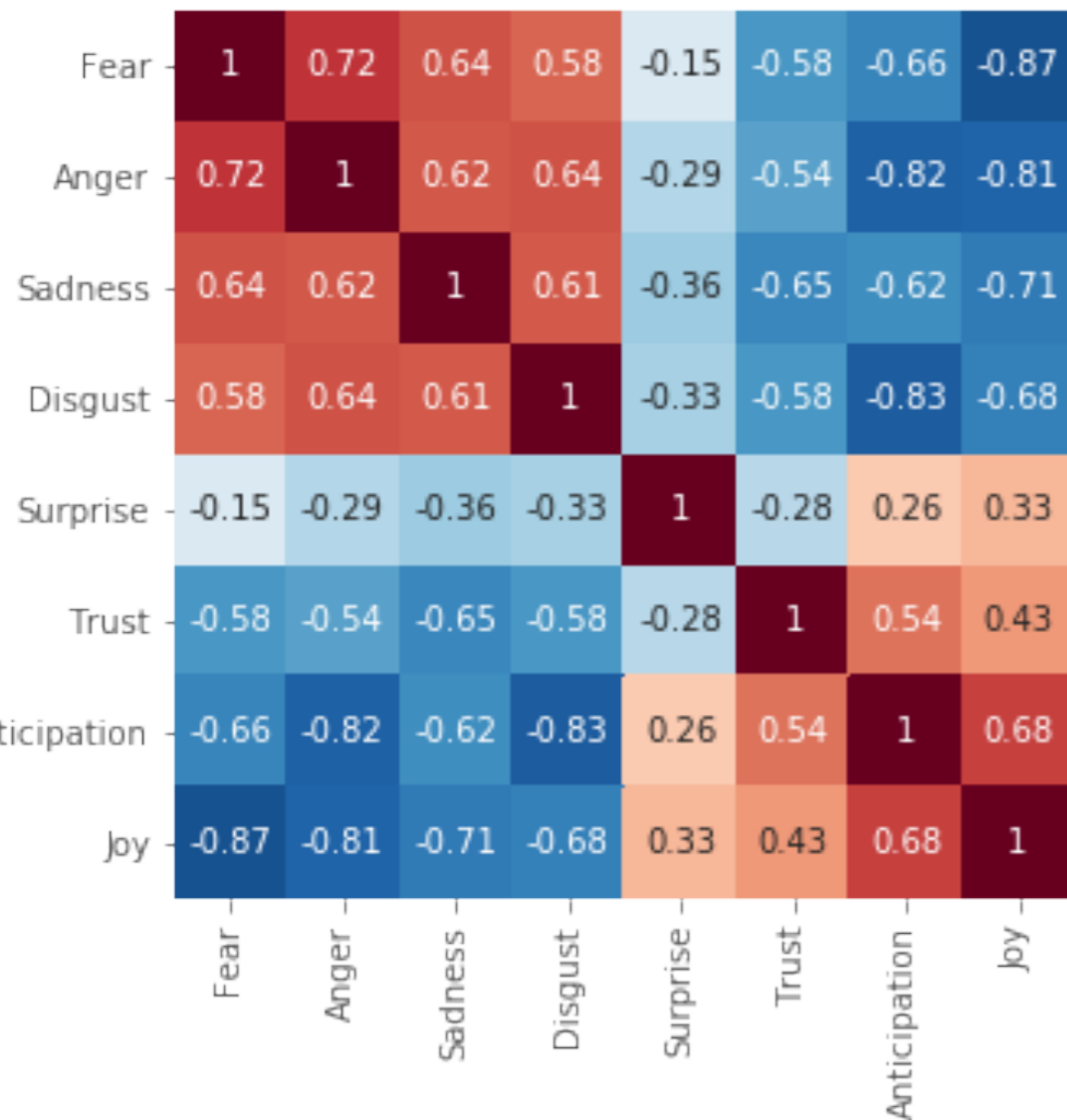
# Correlations

between emotions using Emoji-emotion similarities



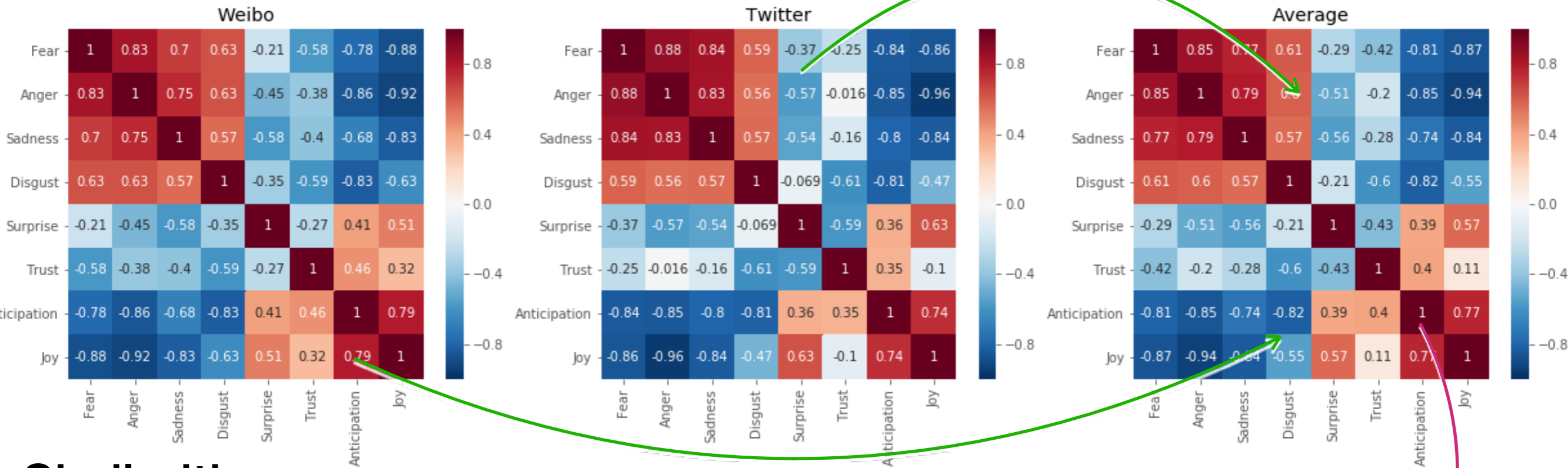
# Similarities

Cosine similarities between emotion vectors

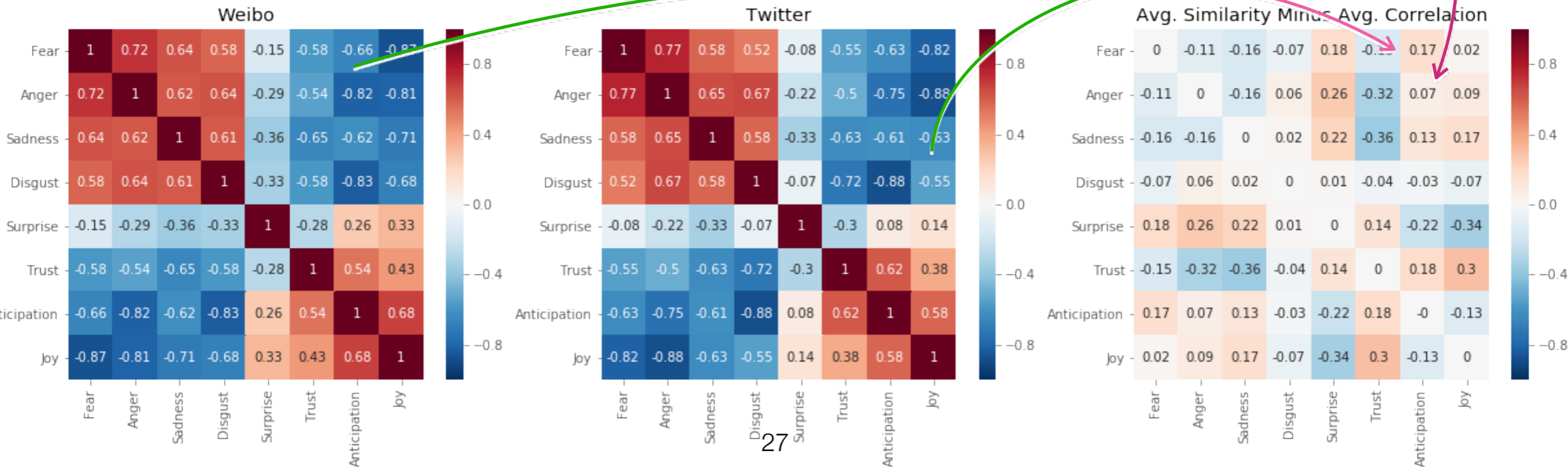


# Put them together

## Correlations



## Similarities



# New Emojis?



Mingyang is a talented comedian! 🌟

















































































the chicken joke was disgusting tho.. 🌟  
Hate it soooo much #notImpressed



# Table of Contents

1. **Two main ingredients:** Emojis and EmoLex
2. **Method:** Corpora, word2vec, average by group, and compare
3. **Results:** Anatomy and panorama
4. **Discussion:** Case study and correlation/similarity
5. **What's Next:** More topics and more emoticons

# Topics, not Emotions

| LIWC             | East  |   |   |   |   | West  |   |   |   |   |
|------------------|---|---|---|---|---|---|---|---|---|---|
| Category         | 1   | 2   | 3   | 4   | 5   | 1   | 2   | 3   | 4   | 5   |
| Ingest           |  80%   |  78%   |  78%   |  78%   |  77%   |  78%   |  76%   |  76%   |  75%   |  75%   |
| Feel             |  65% |  64% |  61% |  60% |  60% |  62% |  62% |  60% |  60% |  60% |
| Family           |  70% |  69% |  68% |  66% |  64% |  68% |  67% |  64% |  64% |  64% |
| Motion           |  67% |  67% |  66% |  66% |  65% |  62% |  61% |  61% |  60% |  60% |
| Space            |  64% |  61% |  60% |  60% |  60% |  56% |  56% |  56% |  55% |  55% |
| Negative Emotion |  61% |  60% |  60% |  59% |  59% |  66% |  66% |  65% |  63% |  63% |
| Sexual           |  60% |  59% |  58% |  57% |  56% |  59% |  58% |  58% |  58% |  58% |
| Money            |  73% |  70% |  69% |  64% |  63% |  68% |  67% |  66% |  66% |  64% |

Meat on the West

Seasonal on the East

Father figure on the East

Cars on the West

Trains on the East

Smoking on the East

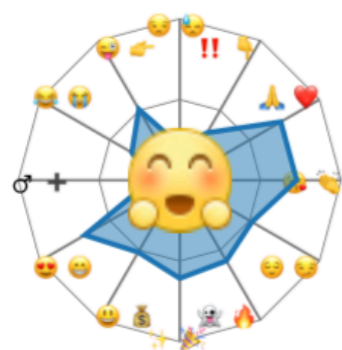
Bathtub on the West

Credit Card on the West

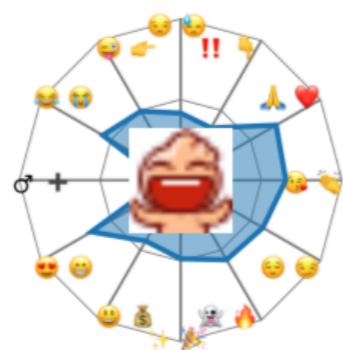
# Map Weibo Emoticons to Emojis, not Emojis to Emotions



Success in success



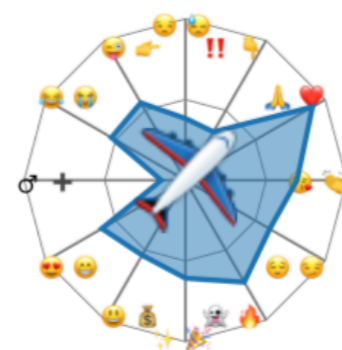
Too happy



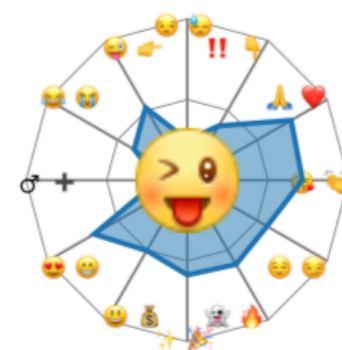
Laughing haha



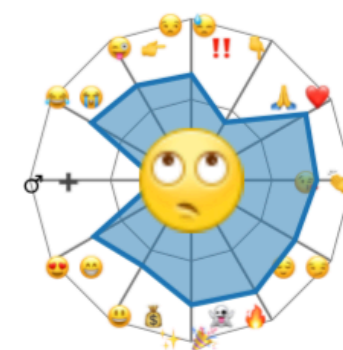
gift



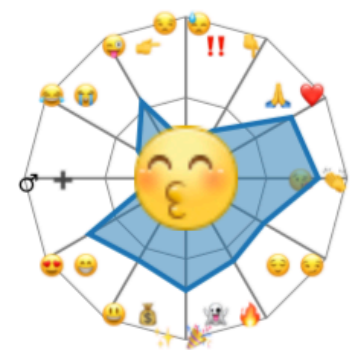
aircraft



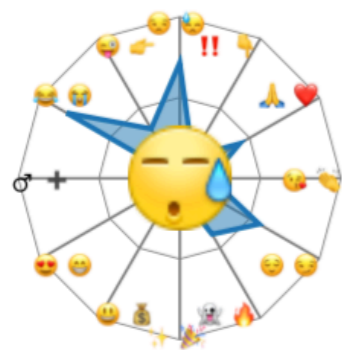
Eye-catching



White eye



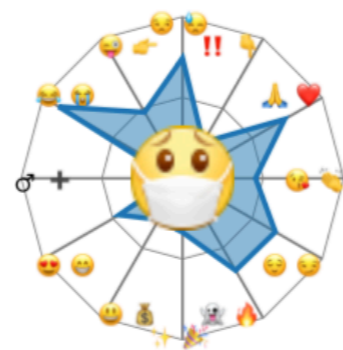
Kiss



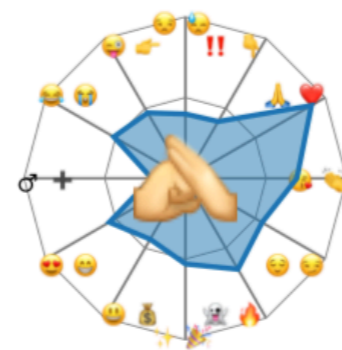
sweat



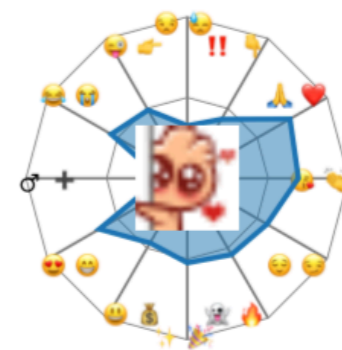
What



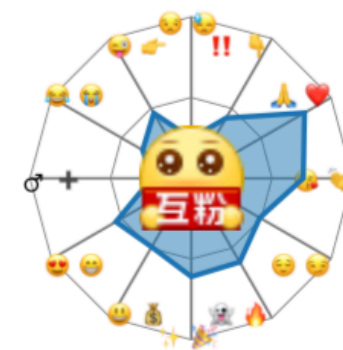
cold



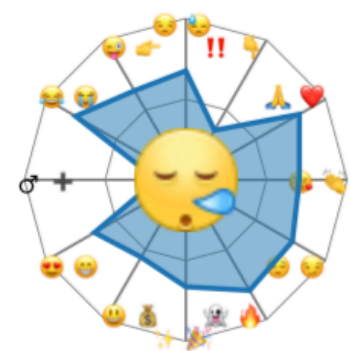
Work



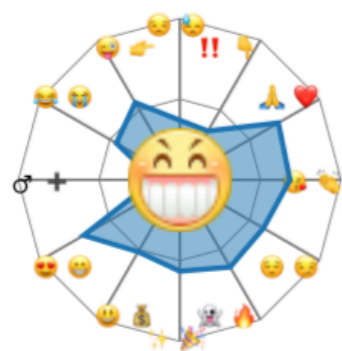
Shy



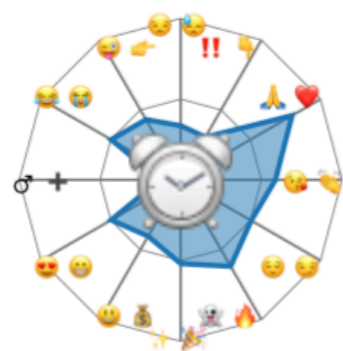
Mutual powder



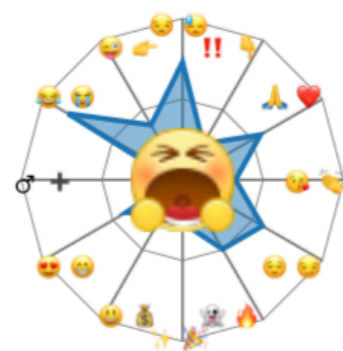
sleep



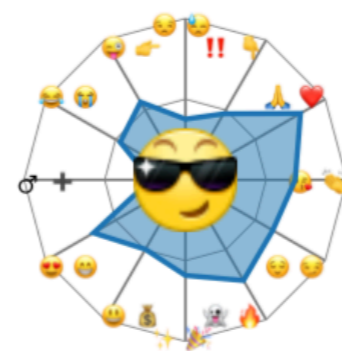
嘻嘻



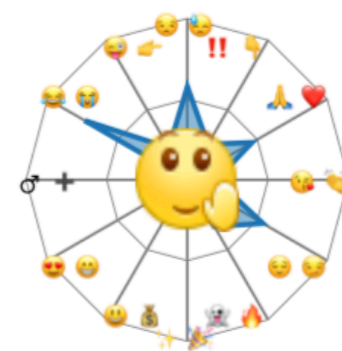
bell



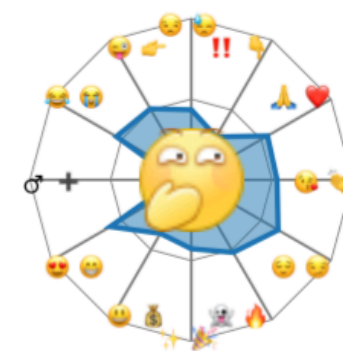
Crazy



Cool



Bye bye



Smirking

# Thank you!

**Special Thanks to my co-authors:**



**Sharath** Chandra Guntuku



**Vinit** Jakhetiya



**Lyle** H. Ungar

*Mingyang Li, University of Pennsylvania*