

# **How regional stereotypes are narrated in news comments: an analysis of the TikTok platform based on the LDA Model**

ExOrdo submission ID: #4288

Presenter: Junyu Jiang

B.A., Communication University of China

E-mail: [jainfarstrider0@gmail.com](mailto:jainfarstrider0@gmail.com)



# CONTENTS

01

## INTRODUCTION

BACKGROUND, PREVIOUS STUDIES, RATIONALE

02

## METHODOLOGY

WHY TIKTOK, SAMPLES SELECTION, BUILDUP OF LDA MODEL

03

## KEY FINDINGS

FORMATION OF TOPICS, SPECIFIC NARRATIVE LINES

04

## DISCUSSIONS AND FUTURE DIRECTIONS





## CHAPTER 1

# INTRODUCTION

BACKGROUND, PREVIOUS STUDIES, RATIONALE

# BACKGROUND

- 01** Regional issues explosive in Chinese social media platform
- 02** Deepening discussion on the image construction and stereotype of Henan (a province of China) people
- 03** Public opinion erupted over the 2023 Nanyang Music Festival in Henan Province and the subsequent theft

# **PREVIOUS STUDIES and rationale**

**Formation of  
Regional  
Stereotypes**

**How it  
Spreads**

**Specials in  
News  
Comments  
Section**



## CHAPTER 2

# METHODOLOGY

WHY TIKTOK, SAMPLES SELECTION, BUILDUP OF LDA  
MODEL

# **WHY CHOOSE TIKTOK AS THE NEW MEDIA PLATFORM**

**User Engagement and Popularity**

**Interactive Features**

**Secondary Comment Areas and Content Production**

# SAMPLES SELECTION

01

## SERACH:

- KEYWORDS: Midi Music Festival, Henan Nanyang, The Midi Music Festival in Nanyang, Nanyang Grab Event.
- 22 effective short video samples were selected

02

## VIDEO ANALYSIS:

- Classify them into four categories: event narration, image restoration, public relations response, and image reflection.

03

## COMMENT DATA COLLECTION:

- Regular expressions in Python and the Jieba word segmentation tool



# BUILDUP OF LDA MODEL

## WHY CHOOSE LDA (Latent Dirichlet allocation)

- An unsupervised learning in text analysis
- Aims to reduce coding errors caused by supervisor bias

## Evaluation and Optimization

- The source text is divided into 1103 bags-of-words models to evaluate topic coherence trends.
- Grid Search is used to optimize the topic model, testing combinations of the number of topics (1 to 14) and learning decay parameters (0.7, 0.8, and 0.9).
- The optimal model is determined by the log likelihood score and perplexity.

## BUILDUP

- LDA training generated 23 topics
- EVALUATION



## CHPATER 3

# KEY FINDINGS

FORMATION OF TOPICS, SPECIFIC NARRATIVE LINES

# TOPIC 1 OBJECTIVE NARRATIVE OF EVENTS

Table 1: Objective Narrative of Events		
TOPIC	FRAME	CHARACTERISTIC WORDS
1	To state objective events.	villagers, theft, belongings, take away, train station, go home, Nanyang, theft, music festival, large-scale, heard, know, believe, mistaken for, underpants, tent, extremely rare, smear
2	To propose recommendations for handling the event.	call for, submit, should, must, do not, commit a crime, sentence, punishment, severe penalty

- 1.1.1 It's probably taken by the nearby villagers.
- 1.1.2 This is definitely theft!
- 1.1.3 It's Nanyang again! The Nanyang Music Festival disaster (This is a pun; the pronunciation of "disaster" in Chinese is like "festival.").
- 1.1.4 The Midi Music Festival didn't make headlines, but the theft did.
- 1.2.1 The government should call on these villagers to collectively return the items.
- 1.2.2 Which village is this? They knowingly committed theft; they must be sentenced.

# TOPIC 2 IRONY NARRATIVE

Table 2: Irony Narrative		
TOPIC	FRAME	CHARACTERISTIC WORDS
1	Objective Statement Form	Procurement, unwanted, respect, customs, people from Henan
2	Allusion in Poetry and Literature	samples above

2.1.1 They even drove cars over; people might think they are stocking up.

2.1.2 What do you mean by stealing? It's called picking up things nobody wants.

2.1.3 Those who have been to Nanyang, please follow local customs. Those who haven't, please respect the local history and customs.

2.2 [197 AD, the second year of Jian'an] During the Three Kingdoms period in the state of Wei, Cao Cao's renowned general Dian Wei had his double halberds stolen by Che Hu'er in Wancheng, ultimately leading to his death in Wancheng (present-day Nanyang).

2.3 Nanyang has the former residences of Zhuge Liang and Confucius. Confucius would ask, what impact and harm does theft cause?

# TOPIC 3 SYMBOLIC DESCRIPTIVE NARRATION

Table 3: Symbolic Descriptive Narrative Logic		
TOPIC	FRAME	CHARACTERISTIC WORDS
1	Distinguishing between events and regional symbols	good person, bad person
2	Distinguishing between Nanyang and the entirety of Henan Province	we the one hundred million people in Henan Province
3	Stating and addressing malignant incidents	Henan, Nanyang, Zhumadian, representative
4	Dialectical form	Y'all betcha, We'uns, Figured

3.1.1 There are good and bad people everywhere; we should condemn the thieves. Regional attacks are unacceptable.

3.2.1 Biased media are also setting the tone. Can Nanyang represent one hundred million people in Henan? Talk about Nanyang, not all of Henan!

3.3.3 Henan has a bad reputation, and the contributions from the Nanyang and Zhumadian (two regionals of Henan Province) area are indeed significant.

3.4.1 Y'all betcha

3.4.2 I reckon nobody wants it.



# TOPIC 4 REFLECTION ON REGIONAL STEREOTYPES

Table 4: Narrative Logic of Reflection on Regional Stereotypes after Image Restoration		
TOPIC	FRAME	CHARACTERISTIC WORDS
1	Exemplifying Outstanding Regional Characteristics	free of charge, well done, do good deeds, friend, proud action, practical action, thankful
2	Proposing Image Restoration Suggestions	Like, contribute, promote, government, volunteers, strict enforcement, approach, announce
3	Support for Image Restoration Outcomes	Increasingly, believe, hope, go for it
4	Subjective Critique and Emotional Output Regarding the Event	Thief, stealing, crazy, criminal activity, rat droppings, disgraceful
5	Expressing Positive Emotions Toward Hometown	Pride, South-to-North Water Diversion Project
6	Aspiring to Change Public Opinion	Believe, gratitude, speak out, smear
7	Calling for Image Restoration	Effort, change, hope

4.1.1 The younger generation has good qualities. The older generation is disgraceful. This morning I saw many private cars in Nanyang offering free rides in the rain to take people back to their hotels.

4.1.2 No matter when, the four provinces will always be together. As a Hebei person, I believe most Henan people are great!

4.2.1 Thumbs up to the young people working hard to maintain the city's image.

4.2.2 Thanks to the editor for promoting Henan, especially my hometown, Xichuan County in Nanyang. You are welcome to visit.

4.3.1 Please don't let any thief go unpunished.

4.3.2 Nanyang needs a good cleanup, making all Henan people lose face.

4.4.1 As a person from Xichuan, Nanyang, the source of the South-to-North Water.

The background features abstract, textured brushstrokes in shades of teal and blue, primarily located in the top-left and bottom-right corners, creating a modern, artistic feel.

## CHAPTER 4

# DISCUSSIONS AND FUTURE DIRECTIONS

# DISCUSSIONS

## Core feature

The spread of regional stereotypes in new media comments represents a digital shift of symbolic violence, marked by portrayals of poverty and backwardness.

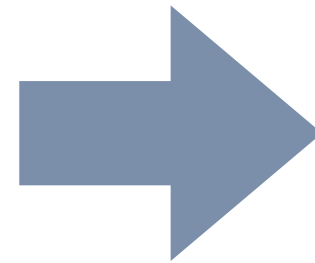
## Media Field and Comment Sections

## Discrimination and Symbolic Violence



# FUTURE DIRECTIONS

Four lines of narrative  
logic  
can be applied to  
**news censorship**



- BALANCED REPORTING
- CULTURAL DISSEMINATION
- POSITIVE IMAGE PROMOTION

## REFERENCE

- Afridi, F., Li, S. X., & Ren, Y. (2015). Social identity and inequality: The impact of China's hukou system. *Journal of Public Economics*, 123, 17-29.
- An, Y. (2020). A history of instructional media, instructional design, and theories. *International Journal of Technology in Education*.  
<https://doi.org/10.13016/5f0da3d8.5f0e2b4c>
- Biot, M. A. (1956). Variational and Lagrangian methods in viscoelasticity.
- Bierstead, R. (1970). *The social order* (3rd ed., pp. 290-292). New York, NY: McGraw-Hill.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Bourdieu, P., Passeron, J. C., Nice, R., Bourdieu, P., & Bottomore, T. B. (2013). *Reproduction in education, society, and culture*. Los Angeles; London; New Delhi; Singapore; Washington, D.C.: Sage.
- Chevtava, E., & Denizci-Guillet, B. (2021). Digital nomads' lifestyles and coworkation. *Journal of Destination Marketing & Management*.
- Deerwester, S. C., Dumais, S. T., Landauer, T. K., Furnas, G. W., & Harshman, R. A. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.
- Duan, Y., Song, X., & Chen, J. (2016). Sense of place: What it means to people. In D. Yifu, S. Xiukui, & C. Jinfeng (Eds.), *Aesthetic Studies: Spring 2016 Volume* (pp. 317-331).
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. T. (2007). Topics in semantic representation. *Psychological Review*, 114(2), 211-244.
- Hu, Q., Pan, X., Luo, J., & Yu, Y. (2022). The effect of service robot occupational gender stereotypes on customers' willingness to use them. *Frontiers in Psychology*, 13, 985501. <https://doi.org/10.3389/fpsyg.2022.985501>

- Hrebícková, M., & Graf, S. (2014). Accuracy of national stereotypes in Central Europe: Outgroups are not better than ingroup in considering personality traits of real people. *European Journal of Personality*, 28, 60-72. <https://doi.org/10.1002/per.1904>
- Huang, L., Clarke, A., Heldsinger, N., & Tian, W. (2019). The communication role of social media in social marketing: A study of the community sustainability knowledge dissemination on LinkedIn and Twitter. *Journal of Marketing Analytics*, 7, 64-75. <https://doi.org/10.1057/s41270-019-00053-8>.
- Jackson-Brown, G. (2010). Understanding community media. *Journalism & Mass Communication Educator*, 65(2), 141-156. <https://doi.org/10.1177/1077695810365957>
- Kaban, A., & Girolami, M. (2003). On an equivalence between PLSI and LDA. In *Annual ACM Conference on Research and Development in Information Retrieval* (pp. 433-434).
- Kravtsov, Y. A., Ostrovsky, L. A., & Stepanov, N. S. (1974). Geometrical optics of inhomogeneous and nonstationary dispersive media.
- Lee, T. L., & Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the Stereotype Content Model. *International Journal of Intercultural Relations*, 30(6), 751–768. <https://doi.org/10.1016/j.ijintrel.2006.06.005>
- Lippman, W. (1965). Public opinion (pp. 63-64). New York, NY: Free Press.
- Lumsden, K., & Morgan, H. M. (n.d.). Cyber-trolling as symbolic violence: Deconstructing gendered abuse online. In N. Lombard (Ed.), *The Routledge Handbook of Gender and Violence*. Routledge.
- Malek, E. (2012). Half-maximal supersymmetry from exceptional field theory. *Fortschritte der Physik*, 60(4-5).
- Marengo, E. A. (2013). A new theory of the generalized optical theorem in anisotropic media. *IEEE Transactions on Antennas and Propagation*, 61(4).
- McNay, L. (1999). Gender, habitus and the field: Pierre Bourdieu and the limits of reflexivity. *Theory, Culture & Society*, 16(1), 95-117.
- Parikka, J. (2016). Deep times of planetary trouble.
- Peng, A. Y. (2021). Amplification of regional discrimination on Chinese news portals: An affective critical discourse analysis. *Convergence*, 27(5), 1343-1359. <https://doi.org/10.1177/1354856520977851>

- Romim, N., Ahmed, M., Talukder, H., & Islam, M. S. (2020). Hate speech detection in the Bengali language: A dataset and its baseline evaluation. arXiv.org [cs.CL].
- Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P. (2004). The author-topic model for authors and documents. In UAI '04: Proceedings of the 20th Conference on Uncertainty and Artificial Intelligence (pp. 487-494).
- Scolari, C. A. (2012). Media ecology: Exploring the metaphor to expand the theory. *Communication Theory*, 5.
- Van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, 4(2), 249-283. <https://doi.org/10.1177/0957926593004002006>
- Vitale, T. (2015). Comment on Loris Caruso/4. Territorial conflicts and new forms of left-wing political organization: from political opportunity structure to structural contexts of opportunities. *Sociologia*, 53(3).
- Urry, J., & Larsen, J. (2011). *The tourist gaze 3.0*. Sage.
- Wetherell, M., & Potter, J. (1992). *Mapping the language of racism: Discourse and the legitimation of exploitation*. Columbia University Press.
- Fu, H. (2010). The paradigm of psychological research on regional culture and regional stereotypes. *Social Psychological Science*, 25(5), 13-17.
- Jiang, Y., & Zhang, H. (2009). A preliminary discussion on cross-cultural research of regional personality stereotypes in China. *Social Science Frontier*, 24(10), 157-160.
- Zhang, H., Jiang, Y., Zhao, W., An, G., Zhang, X., Hu, Z., & Zhang, W. (2012). Theoretical exploration and empirical research on cross-cultural psychology in Chinese regions. *Advances in Psychological Science*, 20(8), 1229-1236.
- Chen, Q. (2022). *Research on stereotypes of Northeast China in rural-themed short videos* [Doctoral dissertation, Changchun University of Technology].
- Wang, X. (2021). The media construction of regional image and the diachronic transmission of stereotypes of Northeast China. *Social Sciences Front*, 2021(02), 267-273.
- Zhang, L. (2016). Regional conflict, attribution of responsibility, and individual attitude from the perspective of stereotypes: An empirical study based on the youth groups of Hong Kong and Guangzhou. *Jinan Journal (Philosophy and Social Sciences Edition)*, 38(05), 59-67+130.