

# Entity Resolution and Location Disambiguation in Ancient Hindu Temples Domain using Web Data

Ayush Maheshwari, Vishwajeet Kumar, Ganesh Ramakrishnan, J. Saketha Nath<sup>1</sup>  
Indian Institute of Technology Bombay, Mumbai, India

<sup>1</sup> Indian Institute of Technology Hyderabad, Hyderabad, India

## Introduction

• Challenge of Entity Resolution in the domain where information is scarcely available. Therefore, data is generated by the crowd over social multi-media platforms such as YouTube, Twitter, etc., Hence, data is ridden with subjective evaluations, opinions, and speculations.

• Data snippets contains multiple temple names and multiple temple location, therefore, additional problem of location disambiguation.

### Irrelevant Postings

The glittering gold-plated gopuram of the Sanctum Sanctorum of the Tirumala Venkateswara Temple is called Ananda Nilayam According to Vaishnava philosophy gopurams of the sanctum sanctorum are named Tirumala Venkateswara Temple Tirumala Ananda Nilayam Sri Ranganathaswamy Temple...

On a february 2017 trip to Sri Lanka we visited 4 World Heritage Site Buddhist Temples. Our guide and tour leader, Janaka Thevin Dissnayake, provided appropriate introductions and commentary throughout.

Palani is a town and a municipality in Dindigul district, Tamil Nadu, located about 100 kilometres north-west of Madurai, 100 kilometres South-east of Coimbatore, and 60 kilometres west of Dindigul

Life in different perspectives from Its tranquil pace will prove jarring for the blockbuster-inclined, but Manakamana rewards patient viewers with a singularly haunting experience- Rotten Tomatoes

### Relevant Postings

Shree Giriniraj Ji Maharaj Daan Ghati Temple is located in Govardhan Mathura UP and is dedicated to Govardhan Lord Krishna. The literal meaning of 'Govardhana' is 'Gov' which translates to 'cows', and 'vardhana' which means 'nourishment'.

Ammachiveedu Muhurthi is a Hindu temple located in the Kollam district of the state of Kerala India It is estimated to be around 600 years old and holds ...

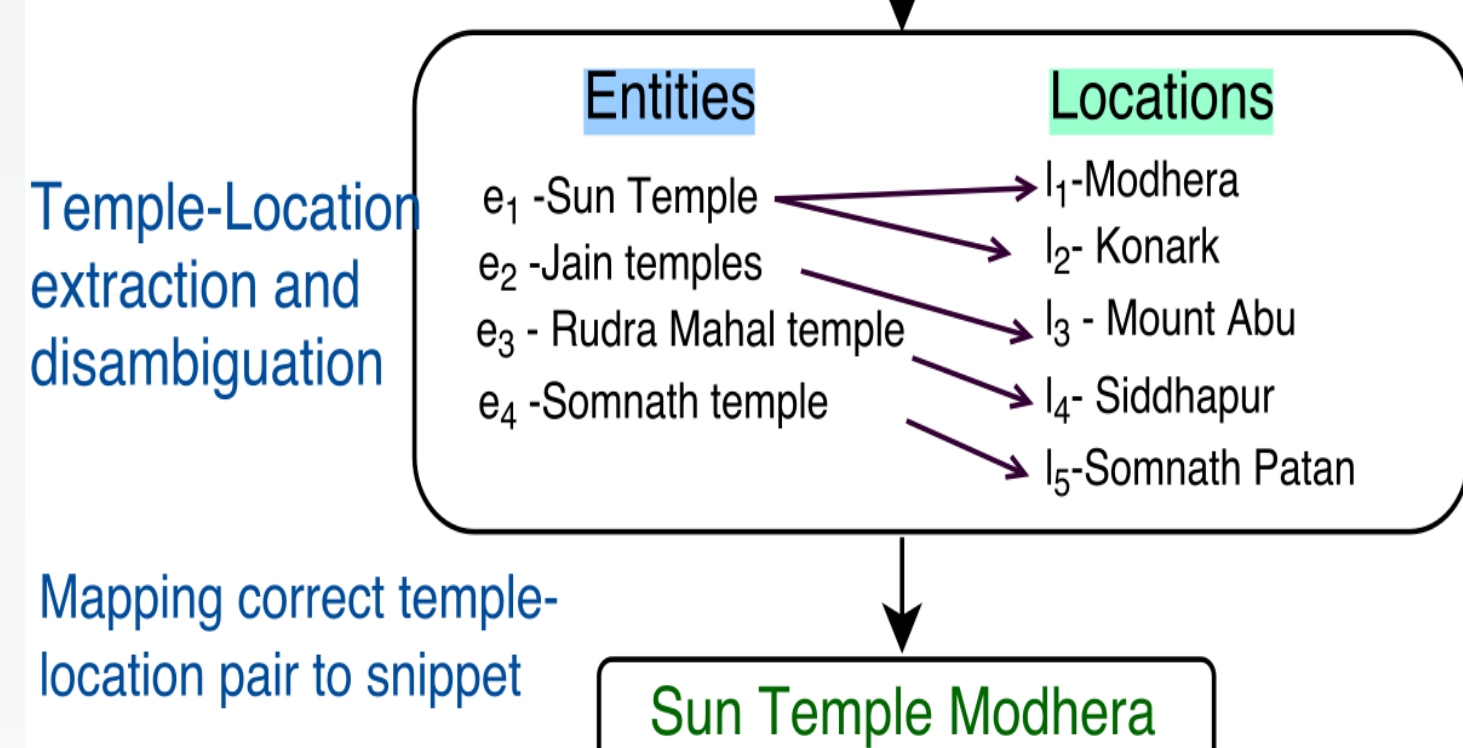
Achalapuram Shivalokathyagar Temple is a Hindu temple located at Achalapuram in Nagapattinam district of Tamil Nadu India ...

Gola Gokaran Nath Temple is a Hindu temple in Gola Gokaranath Uttar Pradesh India and is also known as Shiv Temple of Gola Gokaran Nath

• Sample descriptions (posts) from YouTube videos on various temples. Irrelevant posts are snippets which cannot be associated with any temple, whereas the relevant posts are about Giriraj Dham, Ammachiveedu Muhurthi, Shivalokathyagar and Gorakhnath temples respectively.

## Snippet extraction

The Sun Temple at Modhera dates back to early 11th century CE and was built by King Bhimdev I in 1026 CE. The Sun Temple was built by King Bhimdev I in 1026-27 AD, two centuries before the Sun Temple at Konark. It was during the same period that the Jain temples at Mount Abu were constructed by Vimal Shah. The Rudra Mahal temple at Siddhapur also dates back to this period. Also re-constructed at the same time was the Somnath temple at Somnath Patan.



## Methods

• We manually designed and wrote rules for parsing the textual data using JAPE Grammar in GATE tool.

• Extract temple name and temple location from the document  $d$  using disambiguation methods.

• Use Google Maps API to list temple names located around the extracted temple location. The temple names and temple location form a tuple  $t$  stored in set  $T$

• For each element  $t \in T$ , we calculate TF-IDF score for tuple  $t$  over each document  $d \in D$ , where  $D$  is the indexed set of documents.

• We rank documents based on TF-IDF scores for each query  $t \in T$  and map the top ranked  $d$  to the temple.

## Results

• Our dataset consists of more than four hundred thousand temple names with their locations extracted from Google Places. It also contains more than two hundred thousand videos fetched from YouTube.

• We sample 1000 videos randomly from the complete video set to compute precision, recall and F-measure and evaluate the performance of videos mapped to the temple

Precision = 0.863 , Recall = 0.89 , F = 0.876

## Demonstration

• A demo of our system is accessible at [tinyurl.com/templedemo](http://tinyurl.com/templedemo)

• Demo of the Snippet Ranking system can be accessed at [tinyurl.com/entityr](http://tinyurl.com/entityr)

## References

• Takashi Awamura, Daisuke Kawahara, Eiji Aramaki, Tomohide Shibata, and Sadao Kurohashi. 2015. Location name disambiguation exploiting spatial proximity and temporal consistency. In Proceedings of the Third International Workshop on Natural Language Processing for Social Media, pages 1–9.

• Aliaksei Severyn and Alessandro Moschitti. 2015. Learning to rank short text pairs with convolutional deep neural networks. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 373–382. ACM.

Demonstration of the System

## Methods

• To extract the basic location and video data related to each temple, we use the Google Maps and YouTube API respectively.

• We disambiguate the name and location of each temple using publicly available data on the Web and leverage Google Maps to assign videos to the correct temple.

• Temples names and temple locations are extracted from the snippets using text processing techniques.

• Given a new temple, we retrieve the set of snippets related to the temple. These snippets are fed as input to a CNN based ranking system to score and rank snippets based on the queried temple.