# A Study of Distant Viewing of Ukiyo-e Prints

## Konstantina Liagkou, John Pavlopoulos, Ewa Machotka

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ — ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

STOCKHOLMS UNIVERSITET

## Introduction

- This study targets multicolor landscape prints produced in Japan between the 17th and mid-19th century.

- Defined as *meisho-e* or 'pictures of famous places' these images are grounded in rich poetic culture and tend to mediate depicted places and topographies.

- Due to the richness and diversity of the corpus the extent and nature of this mediation is still not understood by art historians.

- This study hypothesizes that NLP technology allows for a large-scale digital geospatial exploration of the studied prints, which is currently impossible.

- The goal of this study is is to answer the following research questions:
  - What kind of places are depicted in meisho-e prints and what places are not featured in the images?
  - How are these places distributed across Japanese territory?

- The combination of art historical 'close reading' with computational 'distant viewing', or microanalysis of images, is expected to generate a new epistemology of landscape prints.

## Methodology

- NER to identify place-names (LOC & GPE) in inscriptions on ukiyo-e

- The inscriptions on 100 prints were annotated by an art historian

- Used these to fine-tune pre-trained spaCy and BERT models

- Annotated another set of 100 prints for evaluation

- Precision, Recall and F1 scores for evaluation metrics

- The inscriptions of 20 randomly selected prints were annotated by a second annotator, an expert in Japanese history and language

- Computed Cohen kappa for inter-annotator agreement

## Dataset

- Inscriptions on prints are in '**cartouches**' or are positioned directly in compositional space of the image.
- Each inscription includes one or more place-names.
- For example, the inscriptions on Fig. place-name: Tokaido Road. The annotator define 東海道 as GPE.



Figure ... Utagawa Hiroshige (1797-1858), 'Hara' 厚 (*Hara*) from the print series 'The Fifty-three Stations of the Tōkaidō Road: Station Fourteen' 東海道十四五十三次 (*Tōkaidō 14 gojūsan tsugi*), ca. 1847-52, multicolour woodblock print, 22.2 x 34.9cm, The Metropolitan Museum of Art, New York (OA). Inscriptions: Box 1 in the image: 東海道 十四 *Tōkaidō jūyon* (Tōkaidō Station Fourteen). Box 2: 東海道 十四 Mera 渡辺 Watanabe (censor seals). Box 3: 五十三次 *gojūsan tsugi* (The Fifty-three Stations). Box 4: 厚 *Hara* (Hara). Box 5: 広重画 *Hiroshige ga* (designed by Hiroshige).

- Access by Art Research Centre (ARC) at Ritsumeikan University
- 20,408 digitized Japanese early modern ukiyo-e prints
- Our dataset comprises inscriptions of 200 prints
- One main annotator, marked 469 tags in the dataset
- 104 tags were of type LOC and 365 of type GPE



| | Avg | Min | Max |
|---|---|---|---|
| tags per title | 2.33 | 1 | 11 |
| characters per GPE tag | 2.62 | 1 | 8 |
| characters per LOC tag | 2.93 | 2 | 5 |
| cartouches per title | 2.50 | 1 | 11 |
| character tokens per title | 19.68 | 5 | 59 |
| WordPiece tokens per title | 17.17 | 6 | 54 |
| length per token | 1.52 | 1 | 5 |

*Table: Ukiyo-e prints statistics that count of a number of dimensions*



(a) **Characters:** The x-axis depicts the number of characters in the title.

(b) **WordPiece:** The x-axis depicts the number of WordPiece in the title.

## Results

### Inter-annotator Agreement (IAA)
- Micro-averaged Cohen's kappa
- 42.97% for LOC
- 78.63% for GPE
- LOC + GPE → PLACE leads to 78.80% IAA

| | Precision | | Recall | | F1-score | |
|---|---|---|---|---|---|---|
| | SPACY | BERT | SPACY | BERT | SPACY | BERT |
| GPE | 0.84 | 0.73 | 0.44 | 0.74 | 0.41 | 0.74 |
| LOC | 0.59 | 0.59 | 0.50 | 0.50 | 0.54 | 0.54 |

*Table: Evaluation of SpaCy and Bert NER with Precision, Recall and F1-score at the named entity level, on 100 instances. In bold is the best F1-scores.*

### Experiments
- SpaCy and BERT perform equally well in F1 for LOC
- BERT significantly better in GPE
- SpaCy scores high in Precision but lacks in Recall
- BERT balanced between Precision and Recall, with 74% F1

### Error analysis
- Named entities correctly identified
- Misclassification between GPE and LOC, maybe because LOC IAA is low
- Fine-tuning BERT with PLACE labels F1-score increases to 78%
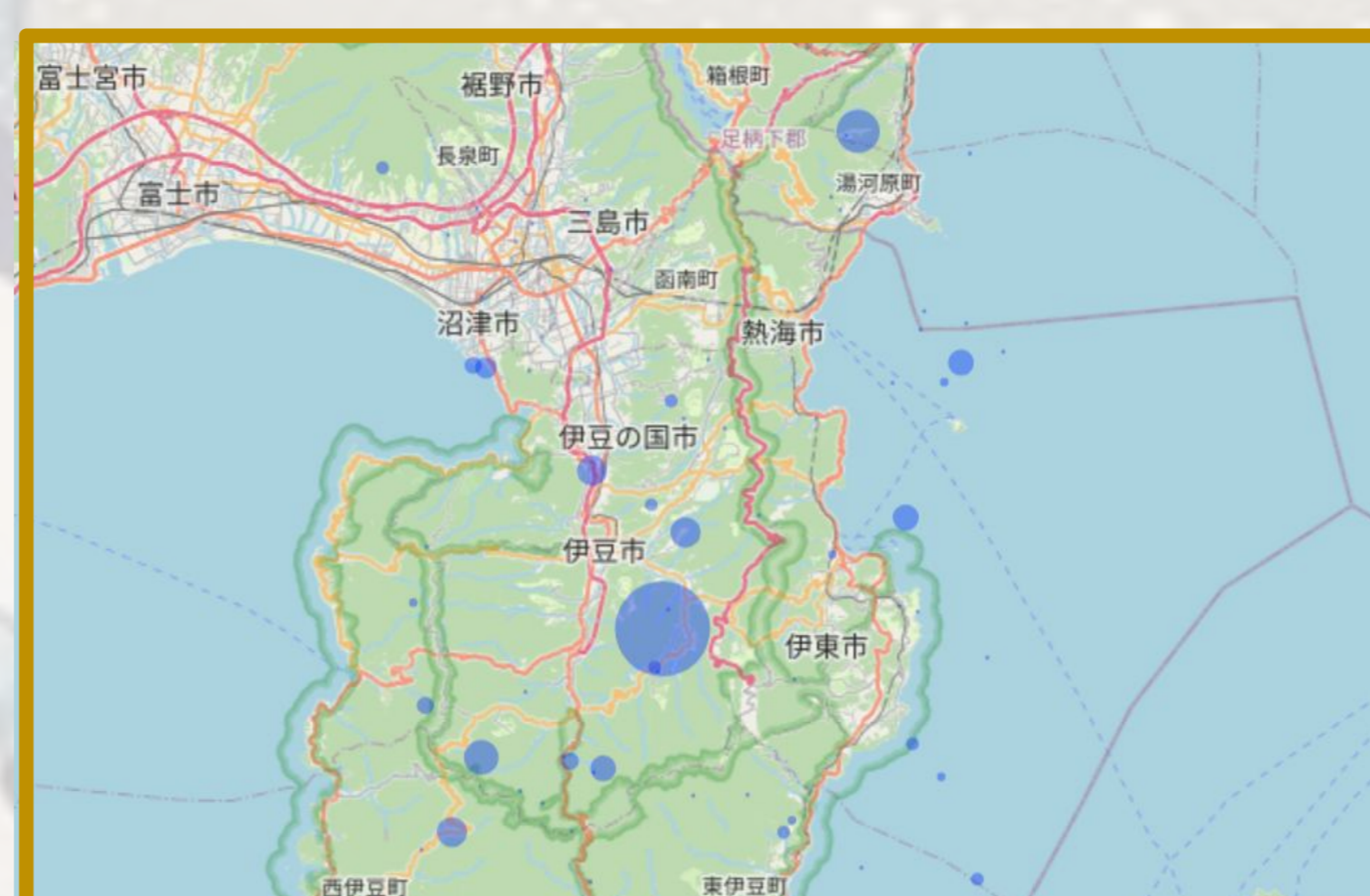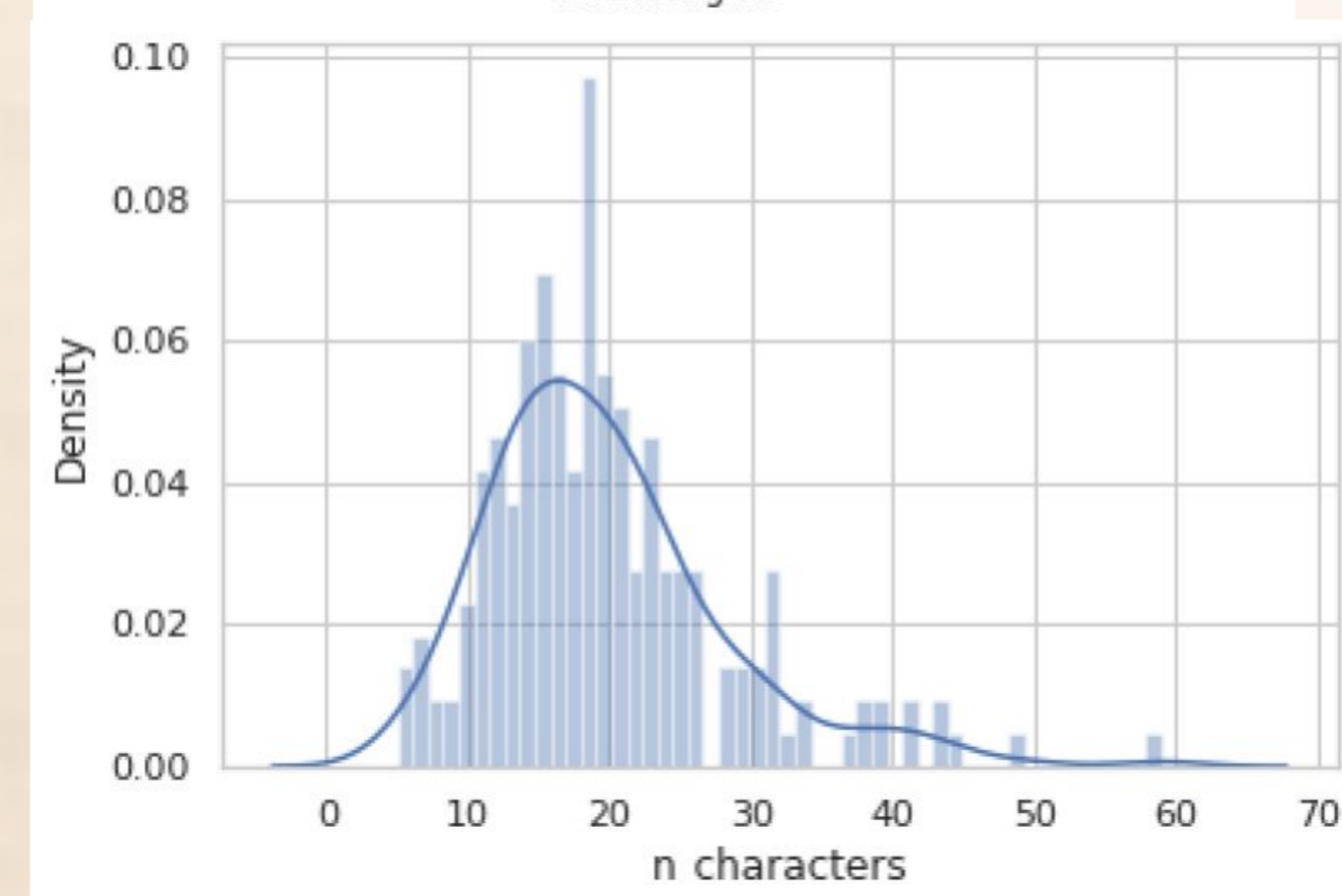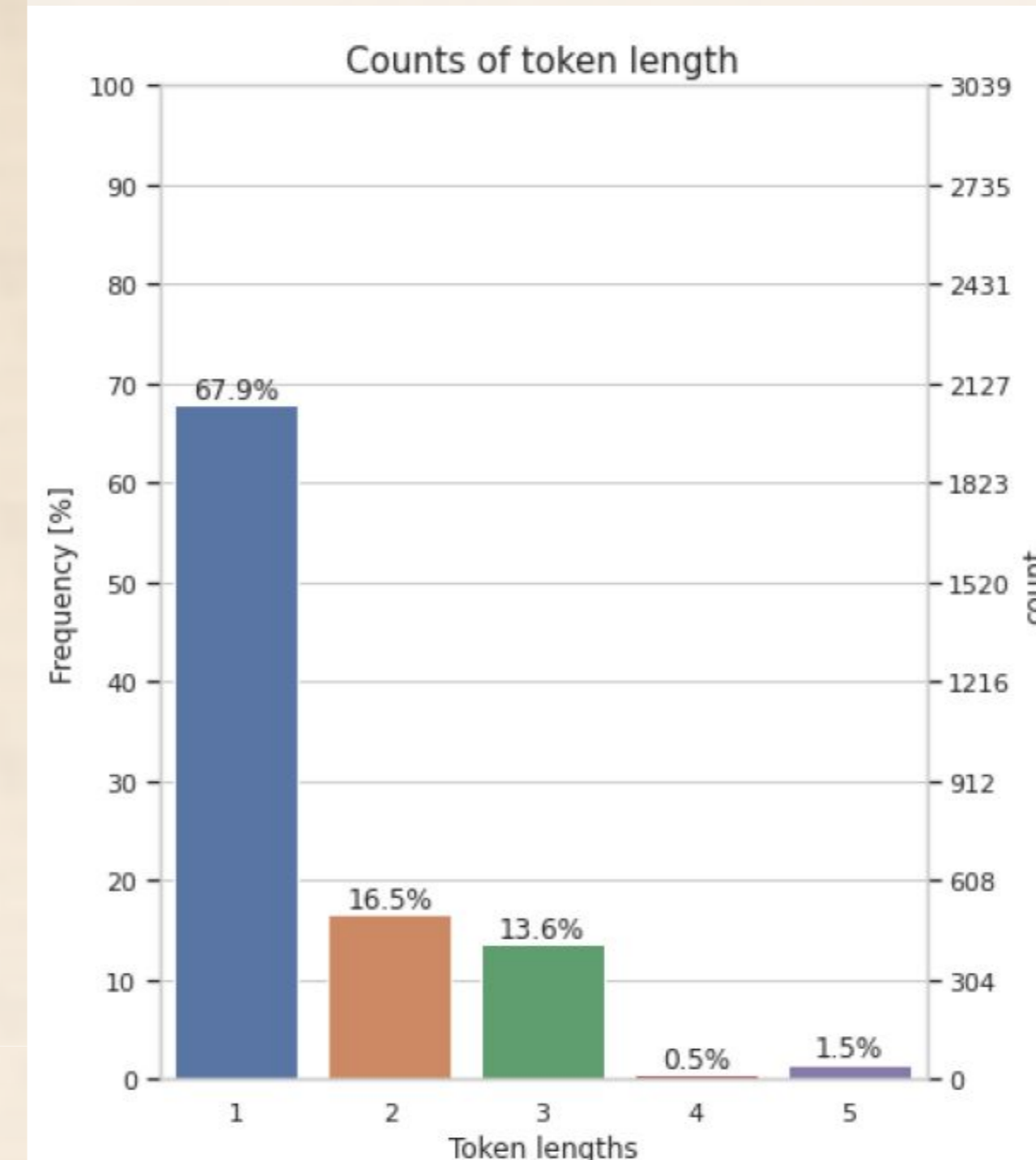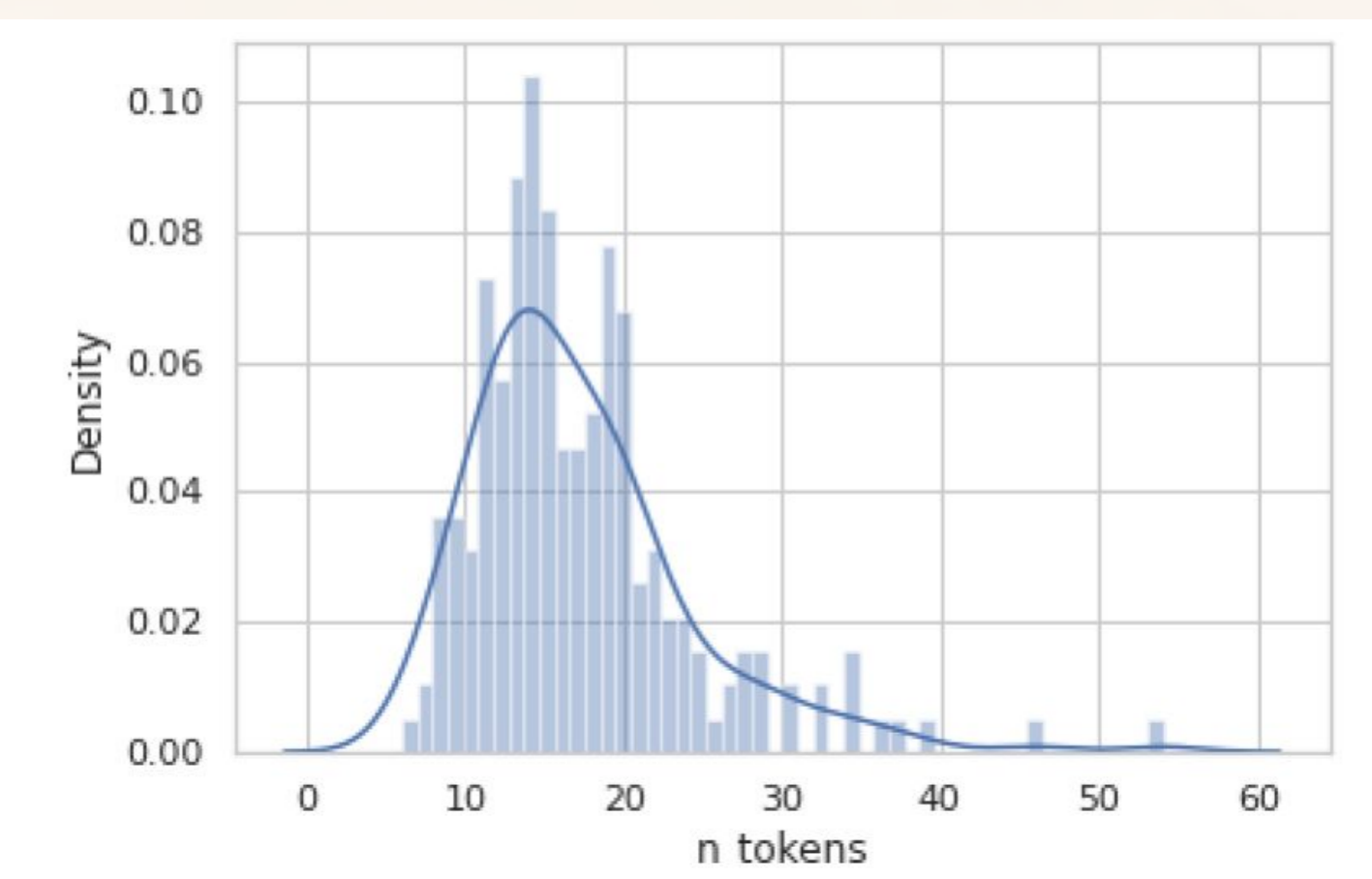
## Discussion

We used the BERT model in order to tag all the place-name entities mentioned within image content-related inscriptions printed on 20,408 prints. The place-names are pinned on a map. The size of each pin reflects the frequency of a given place-name. Tokaido was the most frequent place-name.

| Place-name Entity | Translation | Frequency |
|---|---|---|
| 東海道 | Tokaido | 3901 |
| 江戸 | Edo | 2916 |
| 東都 | Toto | 1609 |
| 東京 | Tokyo | 781 |
| 木曽 | Kiso | 670 |

*Table: The five most frequent place-names recognised by BERT within the the 20,408 titles studied.*



## Limitations

- BERT NER model, fine-tuned on our dataset, can provide a means for 'distant viewing'. However, the model makes mistakes (e.g. places that belong outside Japan).
- Only a single place-name tag, PLACE, was used in BERT NER model.
- The dataset comprises only 200 instances, but more annotations can lead to more accurate models.

## Conclusion

- Presented a dataset (is released for public use) of *ukiyo-e* landscape prints, with place-names included in the print inscriptions annotated by an art historian.
- Japanese BERT-based NER model can achieve a promising performance.
- Use-case of how can a distant viewing of a visual dataset be undertaken for facilitating research in art history.

## Future Work

- Expand our dataset with more inscriptions, as well as with entity types.
- Undertake a spatiotemporal study of ukiyo-e landscape prints and investigate the benefits of NLP-fuelled 'distant viewing' by integrating the dimension of time in our analysis.