Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), pages 5879–5888
Marseille, 20-25 June 2022

A Study of Distant Viewing of *Ukiyo-e* Prints

Konstantina Liagkou, John Pavlopoulos, Ewa Machotka
Athens University of Economics and Business, Greece
Stockholm University, Sweden
konstantinalia4@gmail.com
{ioannis, ewa.machotka}@su.se

Abstract

This paper contributes to studying relationships between Japanese topography and places featured in early modern landscape prints, so-called *ukiyo-e* or 'pictures of the floating world'. The printed inscriptions on these images feature diverse place-names, both man-made and natural formations. However, due to the corpus's richness and diversity, the precise nature of artistic mediation of the depicted places remains little understood. In this paper, we explored a new analytical approach based on the macroanalysis of images facilitated by Natural Language Processing technologies. This paper presents a small dataset with inscriptions on prints that have been annotated by an art historian for included place-name entities. By fine-tuning and applying a Japanese BERT-based Name Entity Recogniser, we provide a use-case of a macroanalysis of a visual dataset that is hosted by the digital database of the Art Research Center at the Ritsumeikan University, Kyoto. Our work studies the relationship between topography and its visual renderings in early modern Japanese *ukiyo-e* landscape prints, demonstrating how an art historian's work can be improved with Natural Language Processing toward distant viewing of visual datasets. We release our dataset and code for public use: `https://github.com/connalia/ukiyo-e_meisho_nlp`.

Keywords: Name Entity Recognition, Natural Language Processing, Digital Humanities

1. Introduction

The recognition of the potential of computational tools and approaches for the study of cultural heritage has been growing in the last decades in different academic disciplines including art history. The advancing digitization of artefacts and growing digital databases as well as the emergence of Open Access principles facilitate and boost these developments worldwide. However, scholars emphasise that the digital research of visual arts must be informed by professional art historical knowledge that acknowledges the situatedness of knowledge production (Bentkowska-Kafel, 2015). Paul Jascot (Jaskot, 2019) poses the question: "(...) what are the critical questions in art history that demand and are best suited to specific digital methods?"

Taking this position as a conceptual point of departure for our research, our project combines 'distant viewing' (Arnold and Tilton, 2019) or macroanalysis of visual materials, and 'close reading' of the artefacts through formal and contextual analysis (Gold and Klein, 2019) to study Japanese early modern woodblock prints, so-called *ukiyo-e* or 'pictures of the floating world' produced between the seventeenth and mid-nineteenth century. These prints rooted in early modern culture of entertainment also include landscape images, which rose to popularity in the first half of the nineteenth century and still attract wide attention worldwide as evidenced by the global career (Guth, 2016) of "The Great Wave" (*Kanagawa-oki nami-ura*, ca. 1830) designed by Katsushika Hokusai (1760-1849). Understood as representations of physical reality, certain places and topography in pre-modern Japan, these prints kept and exhibited by almost all major museums across the world attract large scholarly and public interests. However, it needs to be noted that the artistic genre of 'landscape' is a Western modern art category and only later it was re-applied to pre-modern non-Western objects, which may have performed entirely different social functions at the time of their production. These images were in fact defined as so-called *meisho-e* or pictures of famous places. Importantly, the actual topography of these places was not the guiding principle for their compositions as the concept of *meisho* derives from *utamakura* (lit. poem pillows) poetic rhetorical figures that tie seasonal images with particular places (Kamens, 1997). Therefore, to understand the social function of these images depicting Japanese countryside there is a need to approach them from a different perspective that takes as an analytical point of departure the objects themselves: their content, formal characteristics and socio-cultural and political historical context. As developing an entirely new interpretation of these artefacts is a complex task, it needs to be conducted in several stages starting with a new content-based classification of *meisho-e* guided by the following 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 and in relation to changing historical contexts?
- How does this distribution change in time and depending on the designer, publisher, image format?

The issue is however how to answer these questions considering that the corpus is extremely rich and diversified. For example, the database of Japanese printed culture of the Art Research Centre at Ritsumeikan University, a collaborative partner of this project,[1] alone hosts 678,429 prints kept at 28 institutions in Japan and abroad, and it is not exhaustive. Traditional analytical methods based on close reading or interpretation of selected individual images are not effective for this task. However, the recent development of diverse computational technologies, especially Machine Learning (ML) and Natural Language Processing (NLP) create an exceptional opportunity to engage with this type of inquiry. Motivated by these developments, we employ Named Entity Recognition (NER) to extract named locations from the inscriptions printed on Japanese early modern landscape images. These extracted locations, then, may allow for a large-scale digital geospatial exploration of the studied prints, which is currently impossible. Therefore, in this work, we conduct our analysis and present the first dataset with inscriptions printed on *ukiyo-e* prints, annotated for place-name entities. We release the dataset for public use.

We used this dataset to fine-tune and assess two pretrained Name Entity Recognition (NER) models, vis. spaCy (Honnibal and Montani, 2017) and BERT (Devlin et al., 2018). Our results show that satisfactory performance can be achieved when geopolitical entities and other locations are not separated as entity types. Furthermore, we employ the best performing model as a means of computational 'distant viewing' and address the first two research questions of this study.

2. Related Work

Due to their overall cultural significance in both Japanese and global context Japanese *ukiyo-e* prints have been extensively studied especially since the 1980s that noted the emergence of the so called 'the Edo-boom' rooted, among others, in Japanese economic prosperity and the raise of the discourses on postmodernity conceptually linked to the Edo period (1600-1868) (Karatani, 1989; Gluck, 1998). But despite the booming interest in *ukiyo-e* featuring views of Japanese countryside by the iconic print designers such as Hokusai or Utagawa Hiroshige (1897-1858) the analysis of Japanese 'landscape' images still tends to focus on historiographical, connoisseurial and documentary details (Yamanashi, 2016) rather than larger theoretical questions.

The epistemological implications of the fact that images of the natural environment often took the form of *meisho-e* (pictures of famous places) (Chino, 2003), in which the topography or even locality of particular views was not necessarily relevant (Kamens, 1997) and that the notion of 'landscape' appeared in Japan only in the late nineteenth century (Karatani, 1993) at the apex of the cultural appropriation of Western modernity, have not been comprehensively addressed to date. The research on *meisho-e* prints remains fragmented and often focuses on specific print series or individual designers rather than attempting to look at the visual material at large. This is mainly due to the richness and diversity of the visual material that escape traditional analytical methods based on close reading or interpretation of selected individual images. In this context, computational analytical tools capable of the analysis of large cultural datasets of digitised objects that can facilitate the discovery of new knowledge offer a possibility to rectify this situation.

Despite the recent growth of Digital Art History and access to large digitised image collections, the study of non-Western art images in the context of big data and application of predictive analytics, such as cultural analytics (Manovich, 2020), has been advancing relatively slowly. Different factors are responsible for this situation, ranging from material and financial conditions for research to geo-politics of national identification and international relations. Another important factor stems from the formal characteristics of Japanese pre-modern art and specificities of Japanese pre-modern language, which make computational analysis challenging.

Recent years saw digitization of prints collections and development of print databases worldwide (Akama, 2015), and among others, initiation of computational analysis of *ukiyo-e* prints targeting the questions of style (Tian et al., 2021), attribution (Hirose et al., 2008) and content e.g. images featuring figures of kabuki actors (Yin et al., 2008), developing new tools for automated reading of early-modern inscriptions on prints (Clanuwat et al., 2018; Lamb et al., 2020) or even geolocating of the selected print series depicting Edo city (Suzuki and Kitamoto, 2019). However, these important efforts study relatively small datasets, focus on technical solutions rather than large theoretical questions, and do not target landscape images at a scale and depth that would enable the development of an entirely new epistemology of this art genre. As Natural Language Processing (NLP) can be used to analyse massive amounts of texts, including titles of prints it can assist in the realization of this task. However, despite the expanding role of NLP in a variety of research fields, including Cultural Heritage, its large-scale application for a geospatial analysis of Japanese prints has not been attempted to date.

It is a premise of this study that the combination of 'distant viewing' or macroanalysis of visual materials through the application of NLP technologies, and 'close reading' of the artefacts through formal and contextual analysis facilitated by traditional art historical methods have the potential to develop a more nuanced understanding of Japanese early modern prints, expand and diversify functionalities of digital databases, and in effect, stimulate scholarly and public interests in the non-Western cultural heritage.

[1] http://www.arc.ritsumei.ac.jp/en/index.html

Figure 1: Utagawa Hiroshige (1797-1858), 'Hara' (*Hara*) from the print series 'From the Fifty-three Stations of the Tōkaidō Road' 東海道五十三次之内 (*Tōkaidō gojūsan tsugi no uchi*), ca. 1833-34, multicolour woodblock print, 24.1 x 35.9 cm, the Metropolitan Museum of Art, New York (OA).

Figure 2: Utagawa Hiroshige (1797-1858), 'Hara' 原 (*Hara*) from the print series 'The Fifty-three Stations of the Tōkaidō Road' 東海道五十三次 (*Tōkaidō gojūsan tsugi*), ca. 1840-42, multicolour woodblock print, 15.7 x 20.6 cm, the Metropolitan Museum of Art, New York (OA).

3. New Dataset

First, it needs to be noted that Japanese early-modern prints are not simple snapshots of reality as they are rooted in the culture of popular entertainment related to contemporary kabuki theatre and red-light districts, as well as popular literature, poetry etc. Also, landscape prints, which gained particular popularity in the first half of the nineteenth century, are rooted in premodern textual and visual culture related to *meisho-e* or images of famous places rooted in poetic imaginaries. Therefore, these prints often mediate physical reality. For example, Figures 1 and 2 show two images designed by Utagawa Hiroshige that depict the same location, a small village of Hara known for its breathtaking views of Mt. Fuji and being fourteenth stop on the Tōkaidō road connecting Kyoto, the ancient imperial capital, with Edo (today's Tokyo), the site of the government in early modern period (1600-1868).

The Imagery of Hara Post Station on the Tōkaidō Road

The first image (Fig. 1), printed ca. 1833-34, presents a peaceful view of the countryside with travelers and people working in the fields seemingly unaware of the Mt. Fuji depicted in the background. Its dominating presence is cleverly emphasised by the composition of the image in which the peak of the mountain extends beyond the frame of the picture. The inscriptions printed in the image provide the title of the image and information on its producers. In the upper left corner, the inscription reads the title of the print series 'From the Fifty-three Stations of the Tōkaidō Road' 東海道五十三次之内 (*Tōkaidō gojūsan tsugi no uchi*) and the title of this particular print 'Hara' 原 (*Hara*) followed by a gourd-shaped red cartouche with a white inscription 'Morning Mt. Fuji' 朝の富士 (*Asa no Fuji*). The inscription in the bottom left corner reads 'Designed by Hiroshige' 広重画 (*Hiroshige ga*) and specifies the publisher of this print: Taknouchi Magohachi 竹内 孫八 marked with a red printed seal Takenouchi 竹内.

The shape and proportions of Mt. Fuji depicted in the second, slightly smaller image (Fig 2) is clearly different. Also, a mountain range seen behind a small village depicted in the front of the picture is considerably smaller and differently formed. This image is also entitled 'Hara' 原 and belongs to the print series called 'Fifty Three Stops of the Tōkaidō Road' 東海道五十三次 (*Tōkaidō gojūsan tsugi*), as indicated by the inscription in the red cartouche in the upper right corner of the image. The print was issued by the same designer: Hiroshige (whose signature is found in the bottom right corner) only a couple of years later (between 1840-42). Additionally the image features one more inscription to the left of Mt. Fuji which presents a short 31-syllabe-long poem in the genre of parodic *kyōka* poetry.

Considering the topography of Mt. Fuji (its proportions in relation to its topographical surroundings as well as its shape) viewed from the same vantage point, it is clear that the prints are not faithful representations of physical reality. The images curate reality on at least one more level. They present an idyllic view of the countryside despite the fact that they were produced at the time and directly afterwards of one of the greatest natural, social or political disturbance in Japanese history, so-called the Tenpō Crisis caused by bad weather that destroyed the crops, followed by famine, and widespread socio-political unrest. This paradoxical situation raises the questions of the relationship between reality and its artistic mediation and functions of landscape imagery at the time. In order to answer them, it is necessary to understand the nature and extent to which prints depict and/or distort topog-

Figure 3: Utagawa Hiroshige (1797-1858), 'Hara' 原 (*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: 原 *Hara* (Hara). Box 5: 広重画 *Hiroshige ga* (designed by Hiroshige).

raphy (what is depicted and how) on a larger scale than studying individual images. In this context, computational methods offer a potential to advance art historical inquiry.

Data Access

The access to the data was facilitated by the database hosted at the Art Research Centre (ARC) at Ritsumeikan University, one of the leading Digital Humanities hubs in Japan and a collaborative partner of this project project.[2] The ARC digital databases of Japanese printed culture hosts 678,429 prints kept at 28 institutions in Japan and abroad. Our study investigated 20,408 digitised prints featuring natural environments, issued in the nineteenth century, which corresponds to the final stage of the development of *ukiyo-e* prints, directly before and after the beginning of the modern era in 1868. From the 20,408 digitised prints which arose from our search based on the keywords *meisho* (famous place) and *meisho-e* (image of a famous place), we randomly selected 200 samples to annotate. An art historian, an expert in Japanese early modern history, annotated 200 randomly selected images. The second annotator, and expert on Japanese history and language annotated 20 prints to estimate the task difficulty by measuring inter-annotator agreement. The annotation process was guided by two major principles. First, all places that were possible to be pinned on a map (e.g. names of cities, temples, shrines, bridges) were annotated as GPE. Second, places that were less-easily pinned on a map (e.g. roads, mountain ranges) were annotated as LOC.

3.1. Annotation Process

The annotators worked not on the original inscriptions in the prints but their transcriptions provided as metadata in the database. The database metadata provided not only transcriptions of the inscriptions in the prints but also recorded other textual information printed on the images that is commonly not considered as a part of the image inscriptions. Therefore, these texts are usually not featured in databases and are not researched by art historians. As these inscriptions provide important geographical information, we included them in our analysis, but we note that the analysis of print inscriptions is not without challenges. The issue of the ambiguity of artworks titles, in general, has been recently explored in relation to the computational analysis of Western art (Jain and Krestel, 2019). They noted that it is not uncommon that the artwork titles differ and are

[2] http://www.arc.ritsumei.ac.jp/en/index.html

purely descriptive and only secondarily added to the artefacts by museum curators and art collectors, which makes computational analysis challenging. It is possible to see the same situation in relation to *ukiyo-e* images and this issue is linked to different factors including insufficient linguistic abilities to read pre-modern cursive inscriptions on prints, discrepancies between different collections and database records etc. Importantly, in contrast to earlier *ukiyo-e* prints, landscape prints more than often feature title inscriptions on the image, usually included in colorful cartouches. These inscriptions usually provide the titles of the print series to which the images belong and the titles of individual images, either featured in separate cartouches or not. Besides this, the prints may feature other types of textual inscriptions e.g. names of places depicted in the images, poetry as well as other texts and symbols informing about the producers of the image: its designer, publisher, sometimes other printing technicians involved, and censorship seals. Our current analysis did not include the print production information but focused on the content of the prints, namely place-names featured in titles and other inscriptions. We plan to add more layers of textual information in the future and expand the depth and breadth of our research. Therefore, as mentioned earlier, the annotators worked on the transcriptions of image content-related inscriptions provided as metadata by the ARC database.

The annotators were asked to identify and label two types of place-names. The label GPE was assigned to names of places, both human-made such as temples (e.g., Fushimi Inari Shrine in Kyoto) and natural formations such as lakes (e.g., Lake Biwa), that is possible to geolocate precisely on a map. For example, GPE was assigned to the mentioned names of cities, temples, lakes, etc. The label LOC was given to the names of those places that are not possible to pin on a map with the same precision, such as mountain ranges, rivers, roads, etc. Besides standard place-names, such as 'Kyoto', the annotators were asked to identify a wide range of non-standardised place-names that are not included in The Gazetteer of Japan developed by the Government of Japan and hosted by the Geospatial Information Authority of Japan [3] but have cultural significance rooted in visual and textual culture of Japan. As the Gazetteer of Japan, which includes 3,900 geographical names, follows the resolution of the United Nations Conference on the Standardization of Geographical Names, non-standard place-names are not included. For example, the alternative names for Kyoto (京都), such as *Kyō* (京), *Miyako* (都), or *Kyō no Miyako* (京の都) or *Keishi* (京師) rooted in in pre-modern culture, are not included in the Gazetteer but can appear in the print inscriptions.

The annotators, during the annotation process, referred to a wide range of sources relevant to Japanese visual and textual culture including textual dictionaries as well as prints themselves in order to validate their identifications in relation to the content of the print (what is being represented in them). The exact guidelines that were given to the annotators are provided in the Appendix.

3.2. Inter-annotator Agreement

Two annotators were used to measure inter-annotator agreement for the task. The first annotator was an art historian and an expert in the pre-modern visual culture of Japan. The second annotator was an expert in Japanese history and language. The annotators were asked to identify place-names included in the titles of 20 randomly selected prints. We registered a micro-averaged Cohen's kappa equal to 42.97% for LOC and 78.63% for GPE. When we do not discriminate between LOC and GPE per entity (i.e., all entities are assumed to be of the same type), we report a Cohen's kappa agreement of 78.80%. Hence, we observe that merging the (LOC and GPE) types slightly increases the inter-rater reliability.

Importantly, our inter-rater agreement study indicates that name-entity recognition in inscriptions of Japanese pre-modern prints is not an easy task, not even for experts. To explore this further we focus on the inscription 伏見稲荷お山みち, which one annotator tagged 伏見稲荷 as GPE while the second tagged お山みち as LOC. The discrepancy is due to the specifics of Japanese pre-modern script, which also tend to be used in later artefacts. The inscriptions are often rendered in pre-modern Japanese scripts, used before the standardization of the language in the late nineteenth and twentieth century (Yamaguchi and Frellesvig, 2012). In pre-modern Japanese, the Sino-Japanese kanji characters could be used alternately depending on their phonetic value so the same word could be written in different characters (Yada and Tsutomu, 2012) or they could be written in a phonetic alphabet. Hence, in this inscription みち could mean 'road/path' but there is no kanji character 道 to make it clear. The inscription only features phonetic alphabet, which makes the meaning slightly ambiguous. What is more, お山みち (Oyama-michi) is not a standard geographical place-name to be found in contemporary sources. It is only possible to find 稲荷お山めぐり (Inari Oyama-meguri) which means 'The Tour Around Inari Shrine'. It is not impossible though that there was a path called Oyama-michi around this shrine, and if this is the case the whole title could be annotated as LOC (Fushimi Inari Oyama-michi) or 'The Fushimi Inari Shrine Mountain Path'. But if we assume that we do not know what お山みち (Oyama-michi) means, then it is only possible to identify 伏見稲荷 or 'The Fushimi Inari Shrine' in Kyoto and as such it should be annotated as GPE.[4] These

[3] https://www.gsi.go.jp/ENGLISH/page_e300284.html

[4] The artefact itself is included in the Ritsumeikan Database of *ukiyo-e* prints and is labelled as *meisho* but in fact it is a modern picture that does not represent *ukiyo-e* as

	Avg	Min	Max
tags per inscription	2.33	1	11
characters per GPE tag	2.62	1	8
characters per LOC tag	2.93	2	5
cartouches per inscription	2.50	1	11
characters per inscription	19.68	5	59
word pieces per inscription	17.17	6	54
characters per token	1.52	1	5

Table 1: **Ukiyo-e prints statistics**: the average, minimum, and maximum count of a number of dimensions, from the tags per title, to the token length in characters.

kinds of problems are not uncommon in the process of analysis of other inscriptions on Japanese images.

3.3. Data Exploratory Analysis

The dataset comprises of inscriptions on 200 Japanese early modern *ukiyo-e* prints. This dataset included five prints that represent different editions of the same image. They slightly differ visually (color, grade of wear etc.) but bear the same inscriptions. Our main annotator had marked 469 tags in the dataset, out of which 104 tags were of type LOC and 365 of type GPE. As shown in Table 1, each title comprises two or three tags on average, not less than one and no more than eleven. On average, GPE tags consist of 2.62 and LOC tags of 2.93 Japanese characters. LOC tags range from two to five characters while GPE from one to eight. The image content-related inscriptions on prints that are investigated in this study, are often featured in square boxes called 'cartouches' or are positioned directly in the compositional space of the image (without the box) as can be seen in Fig. 2. Each print may have from one to eleven inscriptions while on average they have 2.5.

The image content-related inscriptions have 19.68 Sino-Japanese kanji characters on average, with minimum 5 and maximum 59. When tokenising to word pieces (see Section 4.2), [5] inscriptions consist of 17 tokens on average, with minimum 4 and maximum 54. The length of the titles, counting in word pieces, follows the normal distribution without a lot of outliers, as can be seen in Figure 4 where very few inscriptions exceed the 50 tokens. The average token length is 1.52, with minimum equal to one and maximum equal to five. Most of the tokens, however, have length equal to one (Fig. 5). In specific, 67.9% of the tokens have length equal to one while approx. 30% of the tokens have a length equal to two or three.

an art genre. However, as it is a part of the database it was randomly selected for our annotation and therefore we decided to discuss it to illustrate the inter-annotator agreement process. This issue exposes the problems with database metadata pertinent to art-historical research.

[5] https://huggingface.co/cl-tohoku/bert-base-japanese

4. Empirical Analysis

This section describes empirical findings, observed by applying Name Entity Recognition (NER) on our data.

4.1. Methods

We fine-tuned the following two pre-trained NER models: spaCy (Honnibal and Montani, 2017) and BERT (Devlin et al., 2018).

SpaCy[6] is an open-source Natural Language Processing (NLP) library that is based on a Convolutional Neural Network (CNN) (Schmidhuber, 2014), pre-trained for part-of-speech (POS) tagging and NER. The default NER model can recognise a wide range of named entities, including place, person, and organization in multiple languages, including Japanese. In this work, we opted for fine-tuning an existing pre-trained Japanese NER model,[7] pre-trained on text from news and media.

BERT has set the state of the art in several NLP tasks and languages. It is a bidirectional transformer (Vaswani et al., 2017), pre-trained on a large corpus using masked language modeling and next sentence prediction. In this work, we opted for a Japanese BERT model,[8] pre-trained on Japanese Wikipedia and fine-tuned with a token classification head on top.

4.2. Experimental Settings

The dataset comprises inscriptions printed on 200 Japanese early modern prints. We used 100 annotated inscriptions for training and 100 for testing. We used SpaCy's tokeniser for the SpaCy model while for BERT we tokenised the texts with a Japanese Tokeniser from Hugging Face.[9]. They were first tokenised by MeCab morphological parser with the International Phonetic Alphabet (IPA),[10] and then split into word pieces (subwords) using the WordPiece algorithm (Schuster and Nakajima, 2012; Song et al., 2020).

A tricky part of NER with BERT is that BERT relies on WordPiece tokenisation rather than word tokenisation. Therefore, we defined the labels at the WordPiece-level. Analytically, we split the inscription into WordPiece tokens, and we shared each tag in all WordPiece tokens contained in the positions of Japanese symbols. For example, a word like '四条河原', which is labeled as 'GPE', it is tokenised to '四条' and '河原'.

For BERT, we used 512 tokens per instance, 256 instances per batch, and 1M training steps. The base layers were initialised with the pretrained weights of "bert-base-japanese".[11] The token classification head of the top had randomly initialised weights, which we trained

[6] https://spacy.io/

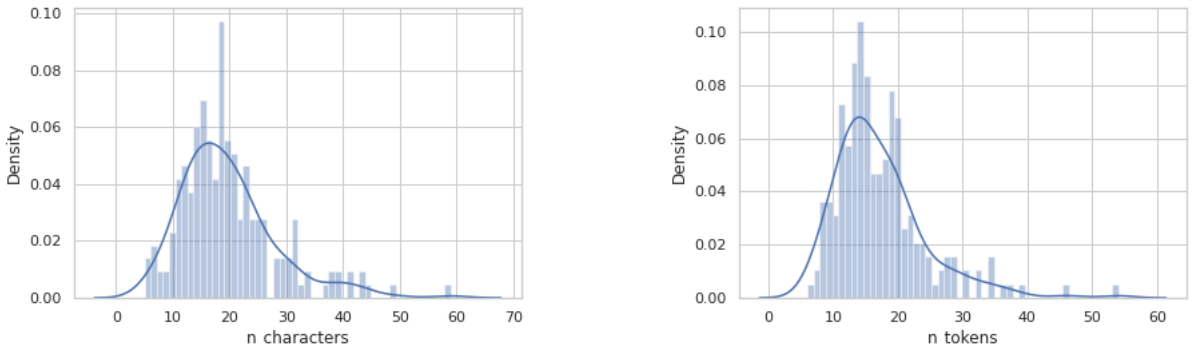
[7] https://spacy.io/models/ja

[8] https://huggingface.co/docs/transformers/model_doc/bert

[9] https://huggingface.co/cl-tohoku/bert-base-japanese

[10] https://taku910.github.io/mecab/

[11] https://huggingface.co/cl-tohoku/bert-base-japanese



(a) **Characters** (b) **Wordpiece tokens**

Figure 4: The density plots of the inscription length follow a normal distribution that is slightly skewed on the right, both when counting in characters (on the left) and when counting in wordpiece tokens (on the right).

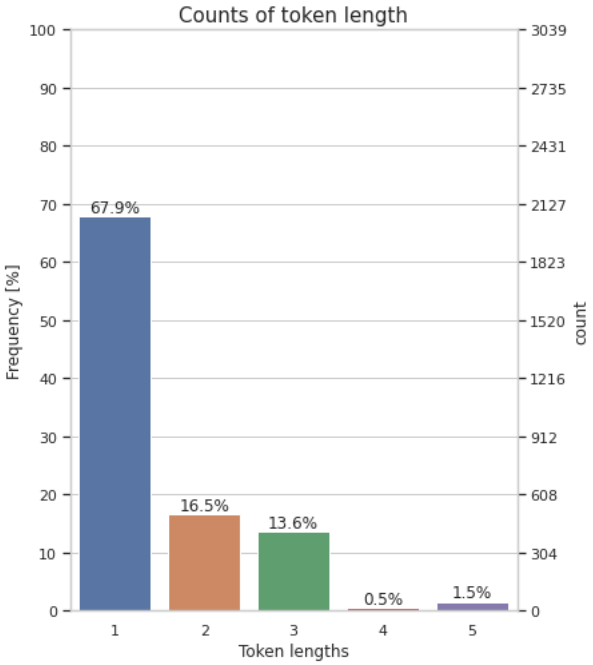


Figure 5: **Token length** measured in word pieces (horizontally), showing the percentage (vertically, left) and the absolute number of tokens (vertically, right). Most of the tokens (over 20,000) have a length of one.

for 20 epochs, together with the pretrained weights, using our dataset. The optimiser was Adam (Kingma and Ba, 2014) with 1e-05 learning rate, and we used a maximum length of 60 to pad the sequences.

4.3. Experiments & Results

Table 2 presents the Precision, Recall and F1 at the named entity level for SpaCy and BERT. The two models perform equally well in F1 for LOC but BERT is significantly better in GPE. Although SpaCy scores high in Precision, it lacks in Recall. BERT, on the other hand, has a better balance between Precision and Recall, achieving an F1 of 74%.

Error analysis showed that the performance of the best performing BERT model is partially explained by its inability to assign the correct labels to the named entities that it correctly identifies. In other words, it can correctly identify named entities, but the assigned labels do not seem to match those of the annotator. This might be due to the fact that annotating Japanese *ukiyo-e* print inscriptions is not an easy task, as it was shown by the limited inter-annotator agreement (Section 3.2). Motivated by our error analysis, which showed that the best performing BERT model confuses the type of place-name in the inscriptions of *ukiyo-e* prints, we experimented with merging GPE and LOC into a single PLACE tag. Then, by fine-tuning BERT on the inscriptions comprising these 331 PLACE labels (258 GPE and 73 LOC tags), we registered further improvements, reaching 78% in F1 (74% in Precision, 81% in Recall).

5. Discussion

We used the best-performing fine-tuned BERT NER model in order to tag all the place-name entities mentioned within image content-related inscriptions printed on 20,408 prints that we retrieved (see Section 3). The place-names are pinned on a map, with the size of each pin reflecting the frequency of the place-name.[12] The most frequent place-names, which are also shown in Table 3, are Tōkaidō, Edo, Tōto, Tokyo and Kiso. This information, which effectively answers the first research question of this work, is only visible through the use of our BERT NER model and it is an example of distant viewing *ukiyo-e* prints.

The landscape prints that were produced at the time, often serialised, belonged to large prints series. These series targeted the most popular topics, such as Edo city or the Tōkaidō Road (incl. 50-100 images). As the inscriptions on these images featured not only names of individual places (e.g., the Hara Station) but also the

[12]The radius r of each pin for a given place-name with frequency f is calculated as: $r = 100 * f^{1/2}$.

	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 2: 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.

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

title of the whole series, certain place-names emerged as the statistically prevalent across the whole dataset. The above mentioned locations, however, refer to the most important places in Japanese pre-modern history, namely Edo (also called Tōto or Tokyo), which was the site of the government, and the two main roads (Tōkaidō and Kisokaidō) that connected it with the rest of the country, including the ancient capital of Kyoto. Hence, the findings are not surprising from the art-historical point of view, but they confirm the interplay between visual arts and the politics of space in Japan.
The map shown in Fig. 6 indicates one more thing. Landscape prints are distributed across different places and they are not only located within these five places. This information comes in response to the second research question, but we note that close-reading should also be employed, in order to provide a more accurate interpretation of these macroanalytical results.

Limitations

- The BERT NER model, fine-tuned on our dataset, can provide a means for 'distant viewing', as shown in Fig. 6. However, the model does make mistakes and noise is expected to be included in this view (e.g., places that belong outside Japan).
- Only a single place-name tag was used by our NER model, because GPE and LOC were merged into a single entity type. However, a bigger dataset could allow an analysis using a much more fine-grained resolution.
- Our released dataset comprises only 200 instances, but more annotations can lead to more accurate models. We plan to extend this dataset with more instances and place-name types.

6. Conclusion

This work presented a dataset of *ukiyo-e* landscape prints, with place-names included in the print inscriptions annotated by an art historian. Our dataset is released for public use. By fine-tuning Japanese Name Entity Recognition models, we showed that a BERT-based model can achieve a promising performance.
We applied our best-performing NER model on a larger dataset of unlabelled prints, providing a use-case of how can a macroanalysis (distant viewing) of a visual dataset be undertaken, as a step towards facilitating the art historian with computational means and specifically with Natural Language Processing. Our analysis showed that the most frequent place-names extracted were the most important places in Japanese pre-modern history, confirming the interplay between visual arts and the politics of space in Japan. Furthermore, by depicting all the extracted places onto a map, we got an indication that landscape prints are distributed across different places and they are not only located within the most important places.
In future work, we aim to expand our dataset with more inscriptions, as well as with entity types. Furthermore, by integrating the dimension of time in our analysis, we will attempt to undertake a spatiotemporal study of *ukiyo-e* landscape prints and investigate the benefits of Natural Language Processing -fuelled distant viewing.

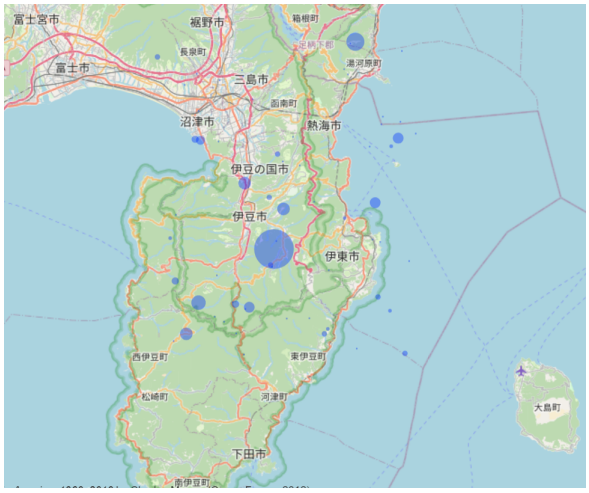


Figure 6: Map with BERT-annotated place-name entities pinned on the map. The area of each pin reflects the frequency of the place.

Acknowledgements

Our work was supported by the Art Research Center of Ritsumeikan University, Kyoto, and partially funded by the Board of Human Science of Stockholm University.

7. Bibliographical References

Akama, R. (2015). 赤間 亮, 立命館大学アート・リサーチセンターの古典籍デジタル化 : Arc国際モデルについて(¡特集¿古典籍資料の最前線), 情報の科学と技術.

Arnold, T. and Tilton, L. (2019). Distant viewing: Analyzing large visual corpora. *Digital Scholarship in the Humanities 34:1: i3–i16.*

Bentkowska-Kafel, A. (2015). Debating digital art history. *International Journal for Digital Art History*.

Chino, K. (2003). The emergence and development of famous place painting as a genre.

Clanuwat, T., Bober-Irizar, M., Kitamoto, A., Lamb, A., Yamamoto, K., and Ha, D. (2018). Deep learning for classical japanese literature. *Neural information processing systems (Neurips) creativity workshop*.

Devlin, J., Chang, M., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding.

Gluck, C. (1998). The invention of edo.

Gold, K. M. and Klein, L. F. (2019). Debates in the digital humanities. *University of Minnesota Press*.

Guth, C. (2016). Hokusai's great wave: Biography of a global icon. *Hawai'i University Press*.

Hirose, S., Yoshimura, M., Hachimura, K., and Akama, R. (2008). Authorship identification of ukiyoe by using rakkan image. *The Eighth IAPR International Workshop on Document Analysis System*.

Honnibal, M. and Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

Jain, N. and Krestel, R. (2019). Who is mona l.? identifying mentions of artworks in historical archives.

Jaskot, P. (2019). Digital art history as the social history of art: towards the disciplinary relevance of digital methods. *Visual Resources*.

Kamens, E. (1997). Utamakura, allusion, and intertextuality in traditional japanese poetry.

Karatani, K. (1989). One spirit, two nineteenth centuries.

Karatani, K. (1993). Origins of modern japanese literature. *Duke University Press*.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization.

Lamb, A., Clanuwat, T., Kitamoto, A., and KuroNet, A. (2020). Regularized residual u-nets for end-to-end kuzushiji character recognition. *SN COMPUT. SCI. 1, 177*.

Manovich, L. (2020). Lev. cultural analytics.

Schmidhuber, J. (2014). Deep learning in neural networks: An overview.

Schuster, M. and Nakajima, K. (2012). Japanese and korean voice searc.

Song, X., Salcianu, A., Song, Y., Dopson, D., and Zhou, D. (2020). Fast wordpiece tokenization.

Suzuki, C. and Kitamoto, A. (2019). Pre- modern japanese books as data of humanities: Finding image of edo famous place from meisho- ki 名所記 and meisho- zue 名所図会 using iiif curation platform. *Keynote Session*.

Tian, Y., Clanuwat, T., Suzuki, C., and Kitamoto, A. (2021). P2 - ukiyo-e analysis and creativity with attribute and geometry annotation. *International Conference on Computational Creativity, Underline Science Inc*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need.

Yada and Tsutomu. (2012). . • (a historical study of japanese writing system). kyūko shoin.

Yamaguchi, T. and Frellesvig, B. (2012). A history of the japanese language.

Yamanashi, T. (2016). Fūkei gakō: Sekai e no kōkan to shinpan. buryukke: Seiunsha. *University of Minnesota Press*.

Yin, X., Xu, W., Akama, R., and Tanaka, H. (2008). A synthesis of 3-d kabuki face from ancient 2-d images using multilevel radial basis function. *The Journal of the Society for Art and Science*.

A. Annotation guidelines

Annotation Instructions: Please annotate the attached dataset comprising 20 digitised images hosted by the Ukiyo-e Portal Database of the Art Research Center of the Ritsumeikan University, Kyoto. The goal is to identify place-names included in the transcription of inscriptions in the prints featured in the selected dataset. Please proceed as follows:

- Open the attached spreadsheet, which comprises the dataset.
- Identify two types of place-names and use two type of labels: GPE for places that are possible to geolocate with precision such as such as cities, temples, lakes etc.; and LOC for places that are not possible to pin on a map with the same precision such as mountain ranges, rivers, roads etc. Use the following labels for the following type of names:

  GPE:

  谷 (*tani*) valley e.g. 黒谷

  塚 (*oka*) mound, hill e.g. 将軍塚

  原 (*hara*) plain, field e.g. 糺河原

  社 (*sha*) Shintō temple e.g. 河合社

  橋 (*hashi*) bridge e.g. 五条橋

  浦 (*ura*) bay e.g. 二見浦

  池 (*ike*) pond e.g. 広沢池

  寺 (*tera*) Buddhist temple e.g. 銀閣寺

  門 (*mon*) gate e.g. 羅生門

LOC:

山 (*yama*) mountain e.g. 吉田山

川 (*kawa*) river e.g.みたらし川

道 (*dō*) road e.g. 東海道

- Please include place-names of cultural significance and do not rely exclusively on The Gazetteer of Japan hosted by the Geospatial Information Authority of Japan (`https://www.gsi.go.jp/ENGLISH/pape_e300284.html`) as it only features standardised contemporary geographical names.
- In your work you can refer to academic and museum sources on pre-modern visual and textual culture that you are most familiar with, and visit the Ukiyo-e Portal Database of the Art Research Center of the Ritsumeikan University, Kyoto for the respective images, if needed.
- Please identify entities that enable the most precise geolocation, for example the title 京都大仏殿大鐘楼(*Kyoto Daibutsuden Dai-shōrō* or 'The Big Bell at the Great Buddha Hall in Kyoto') is annotated as the following place-name '京都大仏殿' (*Kyoto Daibutsuden* or 'The Great Buddha Hall in Kyoto') and labelled as GPE.

Examples are provided with the attached spreadsheet.