Greetings From! Extracting Address Information From 100,000 Historical Picture Postcards

Thomas Smits[4,1,*,†], Wouter Haverals[3,1,2,†], Loren Verreyen[1], Mona Allaert[1] and Mike Kestemont[1,2]

[1]*Antwerp Center for Digital Humanities and Literary Criticism (ACDC), University of Antwerp, Belgium*
[2]*Institute for the Study of Literature in the Low Countries (ISLN), University of Antwerp, Belgium*
[3]*Center for Digital Humanities (CDH), Princeton University, USA*
[4]*Amsterdam School of Historical Studies (ASH), University of Amsterdam, the Netherlands*

Abstract

This paper details the development and validation of computational methods aimed at creating a comprehensive dataset from a vast collection of historical picture postcards.[1] By connecting three distinct locations – the sender's, the recipient's, and the depicted – the medium of the picture postcard has contributed to the formation of extensive spatial networks of information exchange. So far, the analysis of these spatial networks was hampered by the fact that picture postcards are – literally and figuratively – hard to read. Using traditional methods, transcribing and analyzing a sizeable number of postcards would take a lifetime. To address this challenge, this paper presents a pipeline that leverages Computer Vision, Handwritten Text Recognition, and Large Language Models to extract and disambiguate address information from a collection of 102K historical postcards sent from Belgium, France, Germany, Luxembourg, the Netherlands, and the UK. We report a mAP of 0.94 for the CV model, a character error rate of 7.62%, and a successful extraction rate of 419 coordinates from an initial sample set of 500 postcards for the LLM. Overall, our pipeline demonstrates a reliable address information extraction rate for a significant proportion of the postcards in our data (with an average distance difference between the HTR-determined addresses and the Ground Truth text of 36.95km). Deploying our pipeline on a larger scale, we will be able to reconstruct the spatial networks that the medium of the postcard enabled.

Keywords

Historical Postcards, Spatial Networks, Address Information Extraction, Computer Vision, Handwritten Text Recognition, Large Language Models

[1]The dataset associated with this research can be accessed at DOI: 10.5281/zenodo.10005566. It is open for everyone to explore and build upon, provided proper attribution to this paper is given.

CHR 2023: Computational Humanities Research Conference, December 6–8, 2023, Paris, France

*Corresponding author.
† These authors contributed equally.
t.p.smits@uva.nl (T. Smits); wouter.haverals@princeton.edu (W. Haverals); loren.verreyen@uantwerpen.be (L. Verreyen); mona.allaert@uantwerpen.be (M. Allaert); mike.kestemont@uantwerpen.be (M. Kestemont)
https://thomassmits.eu/ (T. Smits); https://whaverals.github.io/ (W. Haverals); https://mikekestemont.github.io/ (M. Kestemont)
0000-0001-8579-824X (T. Smits); 0000-0002-5687-6787 (W. Haverals); 0000-0002-9785-789X (L. Verreyen); 0009-0004-6389-2680 (M. Allaert); 0000-0003-3590-693X (M. Kestemont)

1. Introduction

Since the late nineteenth century, billions of picture postcards have connected people all over the globe. Combining standardized pictures with room for a written message, the postcard is often regarded as one of the earliest forms of mass media that allowed personal communication on a large scale [27, 25]. Produced in print runs of several ten-thousands, they contributed to the forming of persistent visual stereotypes. However, these visual commonplaces were always combined with personal texts: from a short "greetings from", to longer messages scribbled on every inch of available white space. Especially since the so-called Divided Back Period [29], where the front of the card was used for a photograph or illustration, and the recto side for a written message (left), a stamp (top-right), and address (middle-right), postcards became a medium for conveying countless personal micro-narratives of lived experience, that were highly structured and multimodal in nature (see Figure 1).

Next to these characteristics, the specific spatiality of the postcard has been described as one of the medium's defining features [26, 25]. Typically, postcards are sent from one specific location (Place A) to a destination (Place B). In addition to this, they normally depict (and textually relate to) a third location (Place C), as shown in Figure 1. While it may appear that Place A and Place C are necessarily the same, this is not always the case. Essentially, postcards create a triadic connection between the real-world locations of the sender (Place A) and recipient (Place B), and the constructed location portrayed on the front of the card (Place C), which is often idealized and possibly described in the text. A single postcard links these three places for a specific duration: the period between its sending and receiving. When observed on a larger scale, postcards create extensive, complex, and constantly changing spatial networks of information exchange.

The combination of handwritten messages – often scribbled down in varying styles and without much attention to legibility – with images renders the postcard a challenging historical source to decipher and study [8]. As a result, most studies focus on close reading a small number of postcards. Capitalizing on the large-scale digitization of postcards by online auction platforms, this paper presents the first step towards a comprehensive distant reading of the postcard medium. It describes a pipeline that fuses Computer Vision (CV), Handwritten Text Recognition (HTR), and Large Language Models (LLM) methods to extract and disambiguate structured address information from a large collection of handwritten postcards. Although we obtained a dataset of ∼102,000 cards (sent from Belgium, France, the UK, the Netherlands, Germany, and Luxemburg), the present paper presents a pilot study on a representative subset of these as a proof-of-concept. We (1) train a CV model (YOLOv8 [13]) to identify the address regions on the back of the cards, (2) apply a transformer-based HTR model (Transkribus' *Text Titan I* [28]) to convert the identified regions into machine-readable text, and (3) use an LLM (GPT-4 [23]) to extract, disambiguate, and structure address information from these texts. This paper presents results for the CV model (0.94 mAP), the HTR model (7.62 character error rate), and the GPT-4 disambiguation task. For this last task, we propose a simple metric that adequately captures the average distance between the proposed address and the correct address.

Figure 1: A postcard showing the Eiffel Tower (Place C), posted from the Gare du Nord (place A) to the Belgian village of Sint-Amandsberg near Ghent (place B) in 1934.

.

2. Background: postcards and historical spatial information

Despite their omnipresence as a medium of mass communication in the last 150 years, scholarly attention for postcards has been described as 'inconsistent at best' [30]. Historians have mostly viewed postcards as pieces of trivial and insignificant popular culture. As a result, most work on the medium is popularizing in nature and anecdotal in coverage, providing readers with images of bygone times of a specific place or subject [27]. In recent years, the study of picture postcards was reinvigorated by comparing them to (social) digital media, such as text messages, email, and micro-blogging services for image sharing, such as Instagram. Scholars noted that these contemporary 'new media' carry many similar features to the postcard and provoked similar societal responses [27, 22, 8, 33]. Using this analogy, historians have studied how postcards popularized and standardized concepts and knowledge. For example, [20, 38] note how cards played an important role in popular visual nationalism. Pointing to the same underlying process, others have shown that postcards contributed to disseminating and popularizing colonial and orientalist stereotypes [1, 32, 36, 14, 7].

The fact that the postcard is a complex historical source, which is – literally and figuratively – hard to read, might explain the methodological focus of most studies on close reading. For example, [14] uses a sample of only ten cards to examine the 'cross-imperial production and reception of picture postcards from the Dutch East Indies.' Similarly, [2] uses only six cards to draw broad conclusions about the 'voyeuristic economy of the colonial gaze,' which transforms 'other cultures into objects for analysis.' While the meaning of individual postcards might be complex, Pyne [27] points out that, on a larger scale, the meaning of cards is often closely related: 'the more one looks through thousands of postcards [...] the more predictable and samey [they] start to seem'. This paper argues that the specific medial features of the postcard, especially its highly structured multimodal structure, make it a perfect candidate for a distant reading approach. In other words, computational means have the potential to uncover visual, textual, and multimodal patterns in the vast reservoir of historic postcards – our paper hopes to function as a prolegomenon to such an endeavour.

3. Data

Historical postcards are omnipresent in libraries, archives, antique shops, and flea markets. However, in their original analog form, they cannot be studied at scale. In the realm of postcard address recognition, notable advancements have been made in deciphering handwritten and machine-printed texts to enhance mail delivery systems [31, 21]. Furthermore, there are initiatives for analyzing historical postcards through query-by-example word spotting methods [6]. In the last twenty years, several institutions worldwide have started to digitize their collections of picture postcards [16]. However, for the purposes of this paper, most of them are unusable as they only contain unsent cards (without address information). Next to archives and libraries, a large number of postcards have been digitized to be sold via online auction platforms. We rely on Delcampe [url], a large economic stakeholder in this domain, which offers millions of postcards for prices as low as €0.05. Using the website's architecture, we were able to download a maximum of ca. 10,000 images per country/spatial category.[1] To lend geographic focus to our work, we focus here on the postcards that depict places in Belgium and its five neighboring countries: France, the UK, the Netherlands, Germany, and Luxembourg. We collected the maximum number of cards from the general country category and their capital cities: Brussels, Paris, London, Amsterdam, Berlin, and Luxembourg City. Next to the front and verso sides of the cards, we extracted a title/description (provided by the seller), the country/city category (provided by the auction side), and the listed price.

We construct two sample datasets, one to train and validate the CV model and one to validate the performance of the HTR and LLM models. The first dataset contains 1,220 randomly sampled (backsides of) postcards. To provide the model with negative and positive examples, the set contains both cards with and without an address. We manually annotated the address

[1]The site displays a maximum of 10,000 results per search query. While the number changes on a daily basis, Delcampe offers around 60 million cards for sale.

regions using rectangle bounding boxes. To validate the HTR and LLM models, we use a second subset containing the addresses of 500 randomly selected postcards. The address regions have been detected and cropped using the trained CV model. We manually transcribed the addresses and recorded the street address, city and country to which the cards were sent. Using the Google Maps API, we added a geolocation for each card.

4. Methods

Our dataset – which is intrinsically hyper-diverse – presents significant challenges: we have to deal with addresses spanning the entire globe, ranging from the late 19th to the 20th century, and which are inscribed in various languages, characterized by an impressive array of handwriting styles. Furthermore, addresses on these postcards are only semi-structured: some contain detailed information, including the addressee's name, street and house number, postal code, place, region, while others bear minimal instructions for postal services, such as simply a name and a village. For instance, Figure 1 features a postcard sent to Sint-Amandsberg, merely identified as 'near Ghent', without the mentioning of a postal code. In addition, the problem becomes more complex when studying spatial networks that transcend linguistic borders. The names of countries, places, and even streets can be spelled differently in different languages. This is an especially pressing problem for multilingual countries, such as Belgium.

Most historical geocoding studies utilize (fuzzy) string matching between addresses from a historical dataset and entries in historical gazetteers or contemporary databases [5, 17]. However, this technique is highly sensitive to the quality and organization of the address strings in the historical data [5]. Even when historical spatial data is transcribed manually from primary sources – a task requiring significant effort – the resulting entries often contain textual inaccuracies. Misunderstandings may also arise from naming conventions adopted for place name variations, such as 'The Hague', 'Den Haag', 'La Haye', and ''s-Gravenhage', all referring to the same location.

Given the nature of our dataset, conventional (fuzzy) string-matching techniques are of limited relevance. We initially extract the addresses from these diverse handwritten images using the technique of Handwritten Text Recognition (HTR). However, while effective, this approach introduces its own set of problems. Specifically, HTR inevitably introduces textual errors due to the considerable variations in handwriting and language in our data. Therefore, our dataset's suitability for traditional geocoding methods is significantly diminished.

In response to the challenges outlined above, we devise innovative strategies to correctly extract machine-readable addresses that allow for effective geocoding. For this project, we conceived a pipeline consisting of four key stages (see Figure 2), which operate sequentially to provide a holistic solution to the task of address resolution in historic postcards:

1. *Extraction*: We use a **Computer Vision (CV)** model to pinpoint and segment address regions on the digitized postcards' back sides.
2. *Transcription*: These isolated address images are then processed using **Handwritten Text Recognition (HTR)**, converting the handwritten data into a machine-readable format.

3. *Parsing*: After text extraction, we employ a **Large Language Model (LLM)** to systematically structure the raw text into organized address formats.
4. *Resolution*: Finally, we assign geographic coordinates through **geocoding** and validate the accuracy of the extracted addresses.

In stage **1**, we train a state-of-the-art YOLOv8 object detection model on the CV train set.

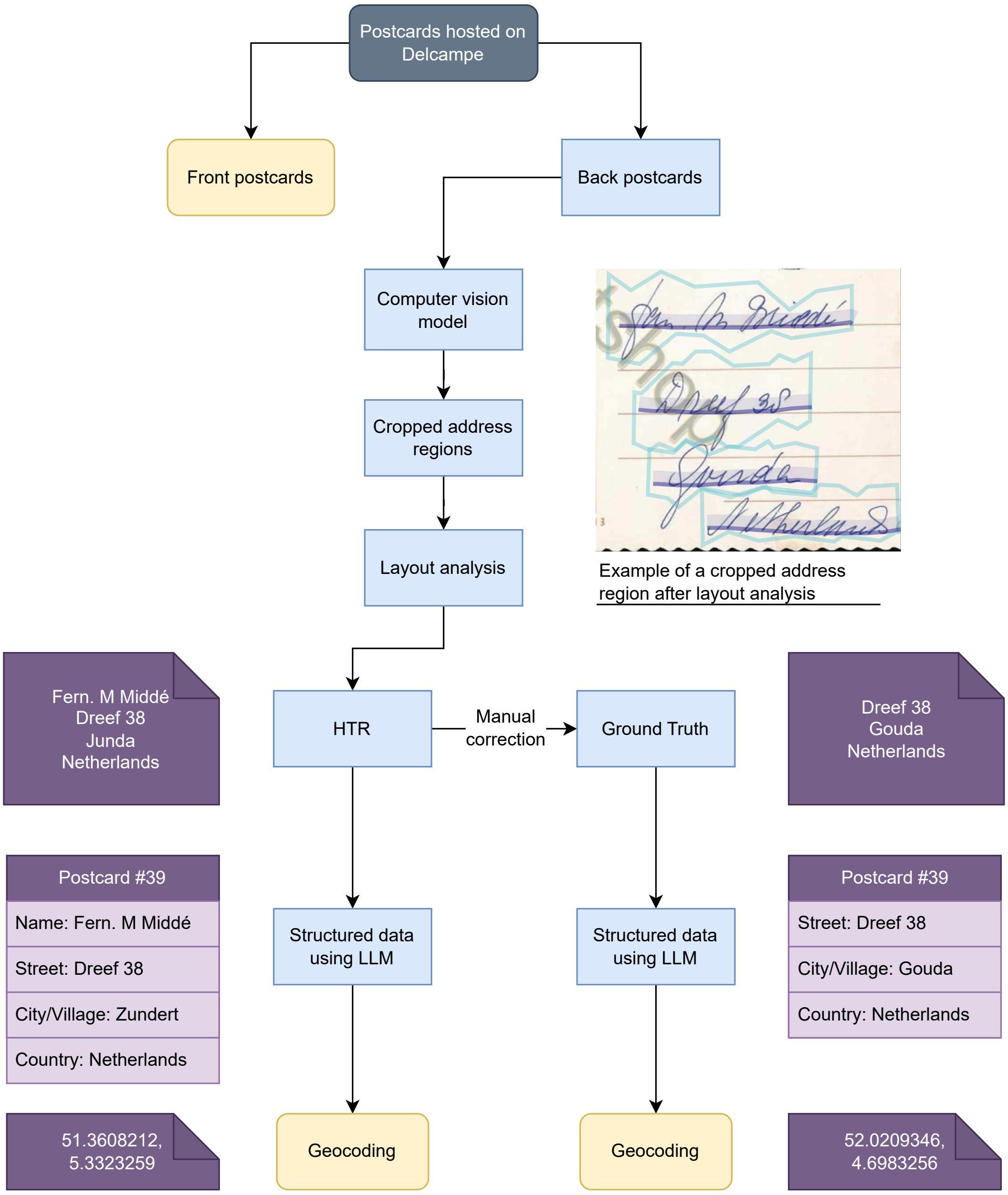

Figure 2: Graphical representation of our pipeline.

Instead of manually selecting hyperparameters, we resorted to the default finetune method and the default parameters (for 30 epochs). To train and validate the model, we use the first subset containing 1,220 randomly sampled postcards, and apply a basic 80/20 split; meaning that the model is trained on a sample of 975 postcards and validated on a second sample of 245 postcards. The model is trained to detect one object class: address region. The randomly selected subset of postcards contains both postcards with and without an address region. The postcards without address are either left blank or contain writing but no address. The majority of the postcards with address have a divided back, with the address region located on the right-hand side. A minority of the postcards with address are undivided, meaning that the back of the postcard only contains the address, written in the center of the card. The trained YOLOv8 model achieves an mAP50 of 0.94 and mAP50-95 of 0.72 on our validation set (predicted bounding box is considered to be correct if it shows an overlap of at least 50% with the ground truth bounding box). After training and validating the object detection model, we use the model to detect and crop the address regions of the postcards used in the downstream tasks.

In stage **2**, we apply the HTR model *Text Titan I* to a random sample of 500 address regions. The address regions are collected by applying the trained YOLOv8 model to our dataset and randomly sampling 500 address regions detected by the model. *Text Titan I*, the recently developed transformer-based 'super model' by Transkribus is one of the most advanced HTR models available today [28]. Given the significant variation within our data set – diversity in image resolution, size, handwriting styles, and language – the decision was made to utilize this robust engine, instead of training our own ad hoc model. *Text Titan I* is particularly suitable for our needs because of its exceptional performance across different handwritings and languages. Using the HTR evaluation package *CERberus*, we observed a Character Error Rate (CER) of 7.62% for our subset of 500 automatically transcribed addresses [11].[2]

In stage **3**, we feed both the automatically transcribed text and the manually corrected text to GPT-4, a Large Language Model (LLM). Following work in several fields that apply prompt-engineering techniques to harness the capabilities of LLM's [10, 35, 15], this was done with two main objectives in mind: (1) to correct potential spelling errors within the addresses, and (2) to impose structure on the raw text. This was achieved by creating a JSON object for each address comprising the following fields: 'Person's Name', 'Street and House Number', 'City/Village Name', 'Postal Code', and 'Country' (where available).[3].

We illustrate the output of our methodology with the example below. The raw text from the HTR model – bearing a spelling error ('Gouda' had been misread as 'Junda') – when fed

[2]This calculation was performed for the address information only. Case-sensitivity, punctuation, and personal names were excluded.

[3]For this purpose, the following prompt was used: "As a sophisticated AI, you're presented with several addresses, each written in multi-line text following the format typically used on postcards. These addresses may comprise a person's name, the street name with house number, the name of a city or village, and occasionally, the country name. However, not all details are consistently provided, and spelling errors may be present. Your task is to identify and rectify these spelling errors, specifically in the city, village, and country names. Cross-reference these details with a comprehensive list of geographical locations. For instance, "Douwersgracht Asteldam" should be corrected to "Brouwersgracht, Amsterdam," and "Brucfel" to "Brussels." Finally, translate this cleaned-up information into a single, uninterrupted structured JSON format. The structure should contain the following fields, if available: 'Person's Name', 'Street and House Number', 'City/Village Name', 'Postal Code', and 'Country'. The purpose of this structured format is to facilitate easier data analysis and ensure uniformity in the dataset."

into GPT-4, results in the LLM model structuring this text into a more organized form.[4] In this process, it modifies ‘Junda’ to a somewhat similar, but incorrect name ‘Zundert’:


```
{"Transcription level": HTR,
"Person's Name": "Fern. M Middé",
"Street and House Number": "Dreef 38",
"City/Village Name": "Zundert",
"Postal Code": "",
"Country": "Netherlands"}
```


In the ‘Ground Truth’ version of this address, the spelling error has been corrected.[5] When we input this corrected text into the LLM, the place name ‘Gouda’ remains unchanged:


```
{"Transcription level": Ground Truth,
"Street and House Number": "Dreef 38",
"City/Village Name": "Gouda",
"Country": "Netherlands"}
```


While it is true that the HTR-generated text holds an error that impacts the derived structured information from the LLM, it is noteworthy that other elements of the address data, such as the country and street name, remain consistent and accurate. Nevertheless, the most significant challenge in our approach arises when the handwriting is challenging for the HTR to interpret, resulting in the introduction of numerous incorrect characters. An example of this issue is when the HTR misreads ‘rue Churchill n 96, Courcelles (Hainaut)’ as ‘Kne Churchill n 96, Camelles (Hamant)’. It is also worth noting that the handwritten addresses can be challenging even for humans to read. We suspect that such difficult readability might even be inherent to our dataset. It is possible that many of the postcards that end up on auction websites were actually left unmailed, due to their hard-to-decipher addresses. Evidence of this lies in the notes on some of the postcards that are marked as ‘Poste restante’.

Feeding both of these raw texts into the LLM, it interprets and structures them as follows. For the manually corrected ground truth text, we get:


```
{"Transcription level": "Ground Truth",
"Street and House Number": "rue Churchill n 96",
"City/Village Name": "Courcelles (Hainaut)"}
```


And for the erroneous HTR-generated text:[6]

[4]It is worth noting that even though the parsing instructions for both sets of text were identical, variations in information structuring emerged. For instance, in the output for the HTR text, an empty ‘postal code’ field is introduced, a feature that is notably absent in the output corresponding to the Ground Truth text.

[5]For the construction of the Ground Truth text, five human annotators looked at the HTR output and suggested improvements. They followed specific conventions during the correction: using ‘#’ for unreadable characters, prefixing lines without address information (e.g., a person’s name) with ‘*’, and prefixing irrelevant lines with ‘@’. Only text pertaining to the geographical address information was corrected.

[6]It is worth highlighting that the LLM, in this scenario, adds a country (France) to the structured output, even though Courcelles is located in Belgium. This not only underscores the occasional unpredictable nature of LLM outputs but also their potential for inaccuracies.

```
{"Transcription level": "HTR",
"Person's Name": "Medames Dennit et Dubois",
"Street and House Number": "Kne Churchill n 96",
"City/Village Name": "Camelles (Hamant)",
"Postal Code": "",
"Country": "France"}
```

The logical sequence of our approach now prompts us to consider the following: how will a geocoder, tasked with translating this structured address information into tangible real-world coordinates, respond?[7] To tackle this, the **fourth and final stage** in our pipeline involves both validating and geocoding these addresses. To accomplish this, we rely on two distinct APIs: the Address Validation API offered by the Google Maps Platform and OpenStreetMap's Nominatim geocoding service [9, 4, 24]. These APIs transform the address data into geographical coordinates, accurately describing their physical locations. Google's API comes with the added advantage of handling potential typing errors, misspelled words, and abbreviations of address elements, efficiently conforming them to both national and international postal address norms. Nevertheless, it also has a downside: its country coverage is somewhat limited, currently only extending to 34 countries. In contrast, Nominatim, while providing support for a substantially broader list of countries and regions, shows little tolerance for spelling errors [12].

Beyond merely obtaining coordinates, we also derive the level of geocoding granularity. This measure serves as an indication of the precision or the level of detail offered by the geocoding process. One of Google's API's unique features is its ability to differentiate between various granularity levels for the interpreted addresses. For our data, both for the HTR addresses and the Ground Truth addresses, we distinguish among the following levels:

- PREMISE: The geocode is accurate up to the level of an individual house or building.
- PREMISE_PROXIMITY: The geocode provides an approximate location at the building-level.
- ROUTE: The geocode offers granularity at the level of a street, road, or highway.
- OTHER: The geocode returned corresponds to a larger area.
- NONE: Both Google's Address Validation API and Nominatim were unable to suggest coordinates.

Our process culminates in this final stage, which also involves quantifying the precision of the suggested coordinates. This step entails determining the average geographic distance, in kilometers, between two sets of coordinates. Each pair consists of one set extracted from the correct address text, and another derived from the text processed by the HTR model. The haversine formula, a mathematical equation frequently employed in navigation, is utilized to perform these calculations. This formula is particularly suitable for determining distances between two points on a sphere using their longitudes and latitudes [34, 19].

[7] We also attempted to request the coordinates from the LLM, but the model hallucinated too often for this to be workable.

To provide a clearer illustration of the final result of our methodology, we present Table 1. This table demonstrates the Ground Truth and HTR-processed structured address information, alongside their associated coordinates and the granularity at which these coordinates are given. The right-most column of the table quantifies the distance in kilometers between these two sets of coordinates, representing the level of precision achieved through the application of our HTR system and geocoding APIs. This distance is calculated using the haversine formula, which provides a reliable measurement of the geographical distance between two sets of coordinates [34, 19]. In doing so, the table also provides insights into the specific discrepancies that arise during the address decoding process.

Table 1
Examples of the output for a handpicked selection of addresses (for illustration purposes) after they were processed by the geocoding APIs.

Ground Truth			**HTR**			**Δ in km**
Structured address	Coordinates	Granularity	Structured address	Coordinates	Granularity	
Herkingen, Holland	51.7102808, 4.0879282	OTHER	*Herkingen, Netherlands*	51.7102808, 4.0879282	OTHER	0.00
Berkelweg 1, 7218 AS Almen, Holland	52.156142, 6.3022203	PREMISE	*Berkelweg, Almen, 7218 AS, Netherlands*	52.1562019, 6.3016427	ROUTE	0.04
Rue Jean l'Aveugle N 7, Arlon, Belgique, Europe	49.6843909, 5.8146424	PREMISE_ PROXIMITY	*Rue Jean l'Aveugle, Liège, Belgium*	50.6560439, 5.5637938	ROUTE	109.51
Dreef 38, Gouda, Netherlands	52.0209346, 4.6983256	PREMISE	*Dreef 38, Zundert, Netherlands*	51.3608212, 5.3323259	PREMISE	85.43
Niška 16/II, Beograd, Jougoslavie	44.8020763, 20.4807197	PREMISE	*Niska 16/II, Belgrade, Yugoslavia*	44.8573492, 20.3783352	PREMISE_ PROXIMITY	10.15
rue du Clair Matin, 71100, St Remy, FRANCE	46.7732417, 4.8305371	ROUTE	*21.100. St Remy., Saoué, France*	44.399109, 2.0396329	OTHER	341.80
######straat 2, Den Haag	52.0704978, 4.3006999	OTHER	*Hendszstraat 2, Den Hage, Netherlands*	52.0866207, 4.3456808	PREMISE	3.56
rue Churchill n 96, Courcelles, (Hainaut)	50.4610782, 4.3851555	PREMISE	*Kne Churchill n 96, Camelles (Hamant), France*	46.227638, 2.213749	OTHER	497.28
58 Rue Ga##d, St Cl###, #####	N/A	NONE	*58 Kur Gounod, S Clone, Deuinctaire*	N/A	NONE	N/A

A review of the examples provided in Table 1 yields several noteworthy results which shed light on both the successes and challenges of our pipeline. One major success of the HTR system and geocoding APIs is demonstrated by their ability to pinpoint accurate geographical coordinates even when slight alterations are made in the structured address, as seen in the cases of “Herkingen, Holland” and “Berkelweg 1, 7218 AS Almen, Holland”. The former produced an identical result, while the latter demonstrated a difference of only 0.04 km. Nonetheless, the table also testifies to the obstacles our method faces. Major discrepancies arise when interpreting addresses with multiple possible interpretations or when important elements of the address are misread by the HTR model. For instance, in the case of “Rue Jean l’Aveugle N 7, Arlon, Belgique, Europe”, the coordinates deviated significantly, resulting in a 109.51km difference, as the LLM that was fed the HTR text misinterpreted the location “Arlon” as “Liège”. A similar issue occurs with “Dreef 38, Gouda, Netherlands” and “rue du Clair Matin, 71100, St Remy, FRANCE”, leading to a substantial distance error. Furthermore, unreadable addresses represented another challenge, as in the case of “58 Rue Ga##d, St Cl###, ######”, which could not be processed and resulted in non-applicable (N/A) outputs. These cases underline the necessity for high-quality text recognition to ensure accurate geocoding results.

5. Results

We present results for all four steps of our pipeline: the CV, the HTR, structuring the data using LLMs, and assigning exact coordinates through geocoding.

1: Identify address regions – Using a small number of training examples, the YOLOv8 model achieves a mAP50 of 0.94, as highlighted in Table 2. As we only train the model to detect a single category (address regions) this high performance was expected. While the mAP50-95 is slightly lower (0.72), we feel confident that the model performs well enough to function in our pipeline. The difference between both metrics can be explained by different standards in how much the bounding boxes of the model and the ground truth should overlap (Intersection over Union). For our task, drawing near-perfect bounding boxes is not of the highest importance and recall should be favoured over precision. After all, most textual information (our focal point of interest) gravitates toward the middle of the box.

2: Automatically transcribe handwritten addresses – Using the general *Text Titan I* HTR model from Transkribus, we report a CER of 7.62% on the address information of the 500 postcards in our dataset. We use CERberus to inspect the CER [11]. This CER is encouraging as a proof of concept, but remains relatively high in comparison to other published work, which is probably caused by the hyper-diversity in the informal handwriting on the cards. However, it is important to emphasize that our dataset essentially boasts as many handwriting styles as there are postcards, a unique challenge that truly puts HTR technology to the test. In this context, only supermodels like *Text Titan I* that are trained on massive corpora encompassing a wealth of variations can handle such a complex task. This highlights the significance of leveraging top-tier HTR models when dealing with data imbued with inherent richness and variety.

3: Disambiguate address information – Our sample subset constituted originally of 500 postcard images. Unfortunately, five of these were of such low resolution that the Handwritten Text Recognition (HTR) model could not recognize any text regions.[8] Consequently, these five cards were omitted from the dataset and all subsequent analyses, leaving us with 495 postcards.

From these 495 postcards, both Ground Truth (GT) and HTR derived text were fed into GPT-4.

[8]Specifically, the problem arises when text regions need to be recognized by the layout analysis model. For these particular 5 postcards, the resolution is too low to recognize any text regions at all.

Table 2
Performance metrics of the CV and HTR tasks in our pipeline.

Task	metric	score
CV	mAP50	0.94
CV	mAP50-95	0.72
HTR	CER	7.62

In some cases, the Language Model did not structure the extracted text as an address but rather treated it as irrelevant text regions. Such content includes messages like "Mit freundlichen, Grüssen", which likely results from too greedy an extraction by the object detection. In these instances, the LLM did not propose an address. This led to 34 Ground Truth texts and 14 HTR derived texts marked by the LLM as void of relevant address information.[9]

It is worth mentioning that the LLM was not prompted to suggest geographic coordinates for the processed addresses immediately. This decision was informed by a preliminary test where the LLM was observed to have a strong propensity to 'hallucinate' by suggesting coordinates that did not match the address information at all. Such hallucinations are a risk at this stage of the method nonetheless (and a danger that has been highlighted in other research as well, see e.g. [18, 39]). An example of this would be when the LLM suggests the country 'France' for French-sounding address text (e.g., because the word "Rue" appears), even when the original postcard does not provide this information. An example of this can be observed in Table 1 where the non-existent place name, but French-sounding HTR text "Camelles (Hamant)" is located in France; while the GT indicates that it is actually a place in the French-speaking Belgian province of Hainaut.

4: Resolution through Geocoding and Validation – In the final step of our method, we assigned geographic coordinates through geocoding and validated the accuracy of the extracted addresses, following the process of address disambiguation. This led us to assess the degree of divergence between the proposed locations for the GT text and the HTR text.

To assess the degree of divergence between the proposed locations for the GT text and the HTR text, two analyses were conducted. Initially, we evaluated the granularity of the suggested geocodes. Figure 3 presents the count of geocodes returned at each granularity level for both GT and HTR extracted text provided to the LLM. Our observations show that the "PREMISE" level has the highest count for GT, while the "OTHER" level tops the count for HTR. This suggests that the manual correction of the geocoded text refines the precision of the address information.

Despite these improvements, there were still instances where place names remained unresolved and did not yield any coordinates from the 495 addresses (70 instances for GT text and 76 for HTR text, as seen in Figure 3). The reasons for this vary. Some texts were not addresses at all but incorrectly recognized text regions on the postcard - 34 instances were noted for the GT text. Additionally, two postcards were found to contain a so-called 'Feldpost' number, a special postcode for items sent via military mail, which cannot be converted into coordinates

[9]The difference in number primarily stems from the GT text being manually checked by human annotators. If a text region was deemed to not contain address information, it was excluded.

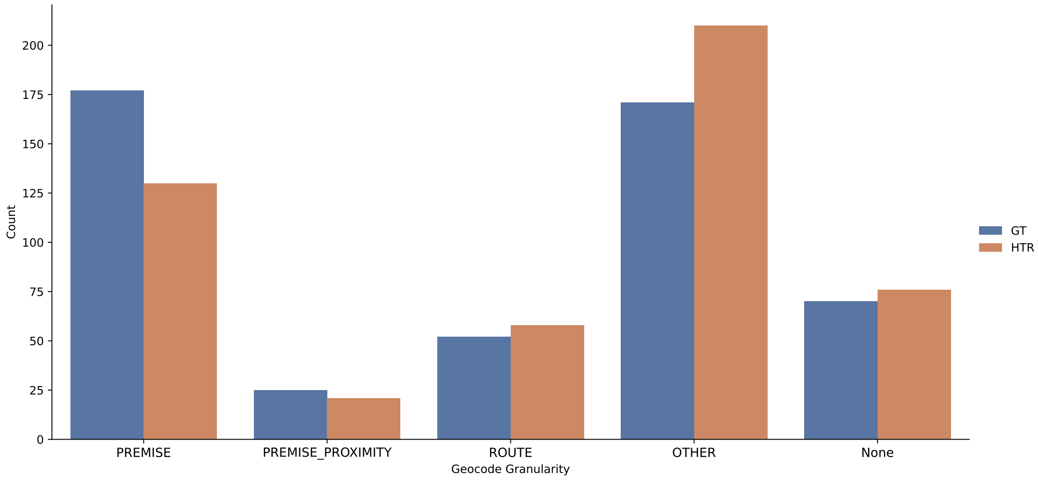

Figure 3: Comparison of geocode granularity distribution for structured address information derived from GT and HTR text.

with our method [3].[10] The remaining texts for which no coordinates could be retrieved by the geocoding APIs were either incomplete, entirely illegible, or simply erroneous addresses. A significant overlap exists between the GT and the HTR text: out of the 70 unlocalizable GT addresses, there were 39 instances where the APIs couldn't suggest a location for the HTR text either. In summary, out of the original 500 postcards, there were 425 suggested coordinates for the GT text and 419 for the HTR text. If we further filter this data to consider only those cases where coordinates were proposed for both the GT and HTR text, we end up with 388 pairs of coordinates. This subset forms the basis for our next stage of analysis: the comparison of distances between the locations suggested by the GT and the HTR methods.

In the subsequent phase, we quantified the distances between the sets of coordinates proposed by the GT and the HTR methods. Out of the 388 comparisons, we obtained an average distance of around 36.95 km (see Table 3). Intriguingly, the median value, along with the 25th and 50th percentiles, register at 0 km. This indicates that more than half of the time, both techniques returned the same set of coordinates. However, the standard deviation of 206.54 km reveals a considerable divergence in certain cases. The maximum distance observed was a sizable 3585.99 km. This extreme result was due to a particularly hard-to-read address. As the human annotator noted "#eg ###, ####, ####", it resulted in the coordinates for "Egypt" (the only legible letters 'eg' forced this interpretation by the geocoding API). On the other hand, the HTR model made an attempt – albeit not very successful – and read "Vig Car, rens Stang", which translated into coordinates for the Danish town 'Vig', that is, indeed, a long way from Egypt.

To better understand these results and go beyond just the numerical summaries, we ultimately constructed a map that can serve as a powerful tool to visually compare and under-

[10]Furthermore, geocoding APIs like that of Google might not always reflect historical geographies or naming conventions, especially concerning places that had their names changed due to colonial rule and subsequent decolonization [37]. Quantifying the extent of this issue poses an additional challenge.

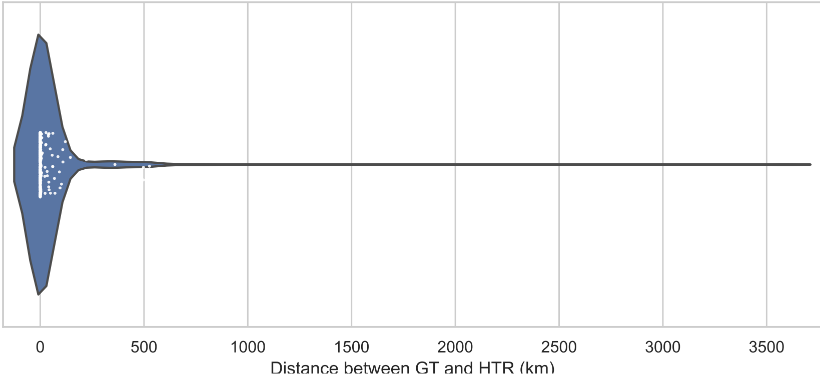

Figure 4: Violin plot of differences in distances between structured GT and HTR address information.

Table 3
Descriptive statistics for the distances between Ground Truth (GT) and Handwritten Text Recognition (HTR) derived geographic coordinates for the 388 postcard addresses.

	Δ in km between GT and HTR
Mean	36.95
Median	0.00
Standard Deviation	206.54
Minimum	0.00
25th Percentile	0.00
50th Percentile	0.00
75th Percentile	0.74
Maximum	3585.99

stand the variations between the GT and HTR coordinates. Figure 5 shows the result of this map, which graphically depicts the geographical locations proposed by both GT and HTR methods, with each method having its own markers. The color of these markers is determined by the distance between the GT and HTR coordinates, with the colormap ranging from dark blue (indicating a smaller distance) to orange (indicating a larger distance). This visual approach allows an intuitive understanding of the geographical spread of the addresses, and more importantly, the variance between the GT and HTR suggested coordinates. A closer inspection of the map highlights areas of low deviation, represented by clusters of blue-colored points. This visual representation supports our initial finding that more than half the time, the two methods returned identical coordinates. However, the scattering of intensely colored points across the map visually emphasizes instances of substantial divergence.

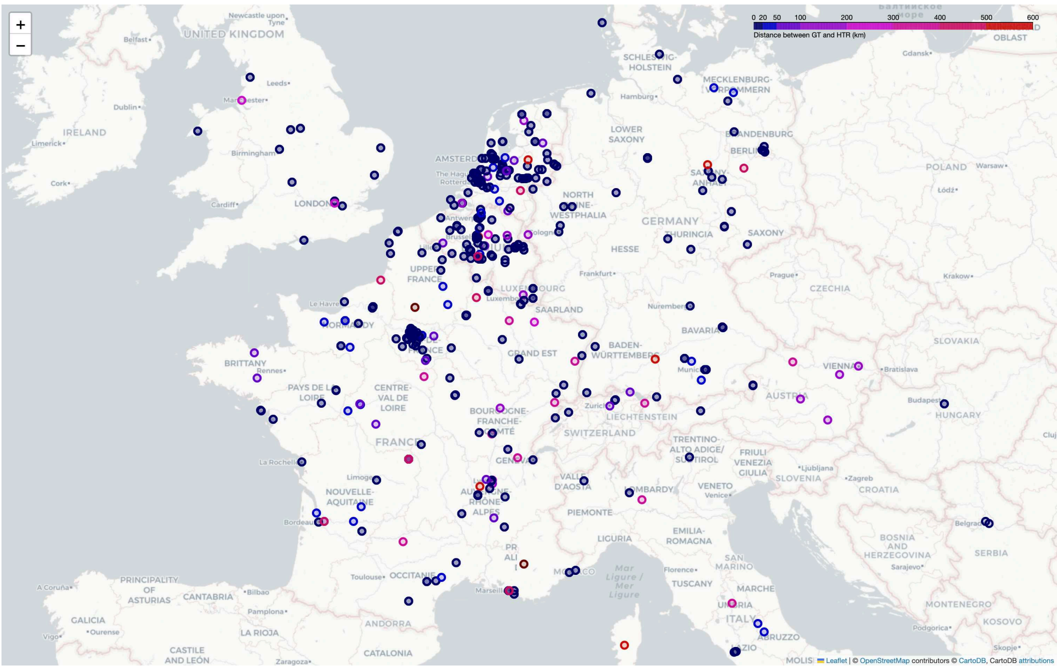

Figure 5: Map displaying GT and HTR coordinates for postcard addresses, with marker colors indicating the distance between corresponding GT and HTR points.

6. Discussion

This paper presented the first step towards a computational distant reading of the postcard medium. In general, we show that our pipeline is effective in extracting spatial information from digitized picture postcards. There are several ways by which we can improve the different steps of our pipeline. For example, the CV model might be improved by providing a larger training set. We achieved notable success with the *Text Titan I* HTR model when dealing with the immense diversity in handwriting. This underscores the necessity and the utility of employing large-scale HTR supermodels for such intricate tasks. Additionally, fine-tuning the prompts might further boost the performance of the GPT-4-based address disambiguation.

An important reflection to make is on the financial scalability and reproducibility of our approach. In our pipeline, we incorporated three commercial products, Transkribus HTR, OpenAI's GPT-4 and Google's Address Validation API. While these offer efficiency and accuracy, they introduce financial implications and potential challenges for widespread reproducibility.[11] To address these challenges, future implementations could explore the use of open-source models or free alternatives that provide similar capabilities.

In future work, we plan to use similar models to extract more and different kinds of infor-

[11]For our dataset of 500 postcards, the total approximate cost was $11.3, composed of charges from Transkribus (5 credits were used, which amounts to approximately $0.8.), OpenAI's GPT-4 (ca. $2 for both prompt and completion), and Google's Address Validation API ($8.5 for 500 postcards). Costs mentioned are based on current pricing as of July 2023. It's noteworthy that these calculations are made without considering potential free tiers or free credits that some services may offer.

mation from digitized postcards. For example, as Figure 1 shows, most sent postcards contain a stamp and a postmark. Combined with the address, these elements can be used to fully reconstruct the journey of the card: where it was sent from (and to), how long this journey took, and how much it cost. In a second avenue of research, we can apply an HTR model to extract the message on the left side of a picture card. Combined with a computational analysis of the pictures on the front of the cards, a distant reading of these texts might tell us a lot about the popularization of specific visual concepts, which can be linked to nationalism, colonialism, Orientalism, and other cultural categories.

While picture postcards have often been dismissed as a trivial or insignificant form of communication, we note that, by approaching them computationally, they offer us the opportunity to discover more about the personal lives of people in the past. In fact, digitized cards offer a vast historic reservoir of untapped micro-spatial narratives of lived experiences. As these personal messages are combined with visual commonplaces, they can also be used to discover more about the connection between personal experience and cultural phenomena, such as nationalism and colonialism. If we are willing to make a trade-off between precision and scale, the presented pipeline offers an interesting instrument for future postcard studies.

References

[1] M. Alloula. *The Colonial Harem.* Vol. 21. Manchester University Press, 1986.

[2] P. M. Burns. "Six Postcards from Arabia: A Visual Discourse of Colonial Travels in the Orient". In: *Tourist Studies* 4.3 (2004), pp. 255–275.

[3] R. Cape. *Youth at War. Feldpost Letters of a German Boy to His Parents, 1943-1945.* New York: Peter Lang Verlag, 2010.

[4] K. Clemens. "Geocoding with openstreetmap data". In: *GEOProcessing 2015* 10 (2015). URL: https://nominatim.org.

[5] R. Cura, B. Dumenieu, N. Abadie, B. Costes, J. Perret, and M. Gribaudi. "Historical Collaborative Geocoding". In: *ISPRS International Journal of Geo-Information* 7.7 (2018), p. 262. DOI: 10.3390/ijgi7070262.

[6] G. A. Fink, L. Rothacker, and R. Grzeszick. "Grouping Historical Postcards Using Query-by-Example Word Spotting". In: *14th International Conference on Frontiers in Handwriting Recognition.* 2014, pp. 470–475. DOI: 10.1109/icfhr.2014.85.

[7] C. M. Geary and V.-L. Webb, eds. *Delivering Views: Distant Cultures in Early Postcards.* Washington D.C.: Smithsonian Institution Scholarly Press, 1998.

[8] J. Gillen. *The Edwardian Picture Postcard as a Communications Revolution: A Literacy Studies Perspective.* Taylor & Francis, 2023.

[9] Google. *Address Validation API.* https://developers.google.com/maps/documentation/address-validation/. 2022.

[10] K. Hatakeyama-Sato, N. Yamane, Y. Igarashi, Y. Nabae, and T. Hayakawa. *Prompt Engineering of GPT-4 for Chemical Research: What Can/Cannot Be Done?* 2023. DOI: 10.26434/chemrxiv-2023-s1x5p.

[11] W. Haverals. *CERberus: guardian against character errors*. Version 1.0. 2023. URL: https://github.com/WHaverals/CERberus.

[12] S. Hoffmann. *Quo Vadis*. 2020. URL: https://nominatim.org/2020/09/14/Quo-Vadis.html.

[13] G. Jocher, A. Chaurasia, and J. Qiu. *Ultralytics YOLOv8*. Version 8.0.0. 2023. URL: https://github.com/ultralytics/ultralytics.

[14] S. Junge. "Familiar Distance: Picture Postcards from Java from a European Perspective, ca. 1880-1930". In: *Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden* 134.3 (2019), pp. 96–121. DOI: 10.18352/bmgn-lchr.10743.

[15] K. Kumar. *Geotechnical Parrot Tales (GPT): Harnessing Large Language Models in Geotechnical Engineering*. 2023. DOI: 10.48550/arXiv.2304.02138. arXiv: 2304.02138 [physics].

[16] M. Ladd. "Access and Use in the Digital Age: A Case Study of a Digital Postcard Collection". In: *New Review of Academic Librarianship* 21.2 (2015), pp. 225–231. DOI: 10.1080/13614533.2015.1031258.

[17] T. Lan and P. Longley. "Geo-Referencing and Mapping 1901 Census Addresses for England and Wales". In: *ISPRS International Journal of Geo-Information* 8.8 (2019), p. 320. DOI: 10.3390/ijgi8080320.

[18] P. Lee, S. Bubeck, and J. Petro. "Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine". In: *New England Journal of Medicine* 388.13 (2023), pp. 1233–1239.

[19] Mapado. *Haversine*. https://github.com/mapado/haversine. 2023.

[20] J. L. Meikle. *Postcard America: Curt Teich and the Imaging of a Nation, 1931-1950*. University of Texas Press, 2016.

[21] V. Mekala, M. Manimegalai, K. Sasipriya, K. Selvakani, and J. Gautham. "Digital Address Identification From Handwritten Address In Postcards". In: *International Journal of Scientific & Technology Research* 9 (2020), pp. 1663–1667.

[22] E. Milne. *Letters, Postcards, Email: Technologies of Presence*. Routledge, 2012.

[23] OpenAI. "GPT-4 Technical Report". In: *ArXiv* abs/2303.08774 (2023).

[24] OpenStreetMap contributors. *Planet dump retrieved from https://planet.osm.org*. 2017. URL: https://www.openstreetmap.org.

[25] J.-O. Östman. "The Postcard as Media". In: *Text & Talk* 24.3 (2004), pp. 423–442.

[26] D. Prochaska and J. Mendelson. *Postcards: Ephemeral Histories of Modernity*. Pennsylvania State University Press, 2010.

[27] L. Pyne. *Postcards: The Rise and Fall of the World's First Social Network*. Reaktion Books, 2021.

[28] Read-coop. *Introducing Transkribus Super Models*. 2023. URL: https://readcoop.eu/introducing-transkribus-super-models-get-access-to-the-text-titan-i/.

[29] B. Rogan. "An Entangled Object: The Picture Postcard as Souvenir and Collectible, Exchange and Ritual Communication". In: *Cultural Analysis* 4.1 (2005), pp. 1–27.

[30] M. Simpson. “Postcard Culture in America: The Traffic in Traffic”. In: *The Oxford History of Popular Print Culture*. Ed. by G. Kelly, J. Raymond, and C. Bold. Oxfored: Oxford University Press, 2011, pp. 169–191.

[31] S. N. Srihari. “Recognition of handwritten and machine-printed text for postal address interpretation”. In: *Pattern Recognition Letters* 14.4 (1993), pp. 291–302. DOI: 10.1016/0167-8655(93)90095-u.

[32] E. R. Stevenson. “Home, Sweet Home: Women and the “Other Space” of Domesticity in Colonial Indian Postcards, ca. 1880–1920”. In: *Visual Anthropology* 26.4 (2013), pp. 298–327. DOI: 10.1080/08949468.2013.804383.

[33] G. Teulié. “Orientalism and the British Picture Postcard Industry: Popularizing the Empire in Victorian and Edwardian Homes”. In: *Cahiers Victoriens et Édouardiens* 89 Spring (2019). DOI: 10.4000/cve.5178.

[34] G. Van Brummelen. *Heavenly mathematics: The forgotten art of spherical trigonometry*. Princeton University Press, 2012.

[35] J. Wang, E. Shi, S. Yu, Z. Wu, C. Ma, H. Dai, Q. Yang, Y. Kang, J. Wu, H. Hu, C. Yue, H. Zhang, Y. Liu, X. Li, B. Ge, D. Zhu, Y. Yuan, D. Shen, T. Liu, and S. Zhang. *Prompt Engineering for Healthcare: Methodologies and Applications*. 2023. DOI: 10.48550/arXiv.2304.14670. arXiv: 2304.14670 [cs].

[36] R. Wehbe. “Seeing Beirut through Colonial Postcards: A Charged Reality”. In: *Department of Architecture and Design: Beirut, Lebanon* (2006).

[37] B. Williamson. “Historical geographies of place naming: Colonial practices and beyond”. In: *Geography Compass* 17.5 (2023), e12687. DOI: 0.1111/gec3.12687.

[38] A. Wilson. “Image wars: the Edwardian Picture Postcard and the Construction of Irish Identity in the early 1900s”. In: *Media Connections between Britain and Ireland*. Routledge, 2022, pp. 30–44.

[39] M. Zhang, O. Press, W. Merrill, A. Liu, and N. A. Smith. “How language model hallucinations can snowball”. In: *arXiv preprint arXiv:2305.13534* (2023).