**An Imagined Geography of Empire: Mapping cultural representations of the Philippines and the American colonial state during the St. Louis World’s Fair of 1904**

Lucas Avelar

Clemson University

December 2023

**Data Ethics Statement**

On the last day of the St. Louis World’s Fair, in December of 1904, Governor Dockery of Missouri stood in front of the Louisiana Monument at the core of the fairgrounds and said that the fair's lesson was made evident to every visiting foreigner: that the United States was "the greatest nation in all the world." He also made sure to address the foreign commissioners gathered around the monument to warn them about an ongoing "war of peaceful conquest." That war, he argued, would prove to every nation in the world the United States' economic and commercial supremacy.[[1]](#footnote-1) The same evening of Dockery’s closing speech, three dead Igorrote bodies were shipped back from Missouri to the Philippines. The Igorrotes, an ethnic group native of the Cordillera Mountain Range in the northern Philippines, had been objectified and exposed at the fair as examples of uncivilized culture and died of pneumonia shortly after they arrived in St Louis in the spring. Their bodies were kept on the fairgrounds for months before fair officials cared to arrange their shipment back home, which finally happened on the last day of the fair. Other sixty-nine Igorrote people were sent back to their homeland that evening along with those dead bodies.[[2]](#footnote-2) In the most extreme level, this tragic case encapsulates how the rhetoric of empire and American colonialism embedded on the fairgrounds depended on the objectification of different ethnic groups and geopolitical entities – in particular of Native Filipinos – to emphasize the exceptional modern identity of the United States.

As a digital history project that deals with the historical experiences of people who have been subjugated, otherized, exposed, and objectified at the World’s Fair of 1904 in St. Louis, MO, *An Imagined Geography of Empire* strives to recognize those individuals with dignity, respect, and ethical responsibility. In the age of “big data” and indiscriminate access and distribution of personal information across profit-oriented digital platforms, historians have an ethical obligation and moral responsibility to acknowledge the lived experiences, memories, and cultural identities of represented historical subjects and/or contemporary stakeholders that are associated with, affected by, or interested in a certain matter of the historian’s work. In this sense, this project is fundamentally informed by the understanding that human interventions to the data and the algorithmic process are a crucial component to every humanistic inquiry that, like *An Imagined Geography of Empire*, uses computational tools for collection, analysis, interpretation, and/or communication.

The data in this project derived from newspaper clippings retrieved from the digital database *Newspapers.com,* the largest online newspapers archive to date. The clippings were stored across 461 JPG files that were then OCR'ed and processed as plain text data in RStudio. The collection was done through both random and proportional sampling, which means that the textual data is proportionally distributed across the three newspapers selected for collection (*The St Louis Republic, St. Louis Post-Dispatch,* and *St. Louis Globe-Democrat*). The data is also proportionally distributed across the seven months of the fair using a fix interval. Yet, the articles chosen to populate each date were randomly picked from the "world's fair" term query results. With a raw count of 196,336 words, the final corpus used for the analysis contained a significant amount of depreciating terms that referred to Filipino people and native groups from North America, Africa, Latin America, and Asia as racially and culturally inferior to white Americans.

Because of the metaphoric nature of the world’s fair as a microcosm of power relations, contradictions, and geopolitical tensions, some methodological choices, and interventions on multiple levels of the research process – the data, the algorithm, the analysis, and the rhetoric strategy for visualizing information from the results – were deemed necessary for the feasibility of the project. First, during the process of named entity recognition and extraction, I considered every mention of nationality and its variations in the corpus (including any OCR errors I could catch) and lumped them together in arbitrarily assigned placenames. So, for example, occurrences of “Brazilian” were lumped together with occurrences of “Brazil” since, even if the mention was referring to the Brazilian pavilion, or the Brazilian coffee, or the Brazilian committee, those mentions were ultimately referring to Brazil as a place in the globe, and therefore contributed to the symbolic production of space and the shaping of an imagined geography of the American empire.

On the fairgrounds, the material culture in display served as metaphors for the represented geopolitical entities, and every mention of nationality, culture, or ethnicity is intrinsically associated with a geographical place in the globe. For this reason, the varied ways in which cultural commentators reflected on and talked about the fair generated discursive representations of the participating geopolitical entities. Those representations are reflected not simply in the placenames printed on the pages, but also in every mention of national and cultural identities and material culture on the grounds that ultimately referred, as a metaphor in itself, to the geopolitical entity represented on the grounds.

Of course, this approach comes with complications. Besides the extensive close reading and “data preprocessing” labor that it implicated, the approach was also faced with the emergence of dubious terminology like “Indian”, which could either refer to the nationality of India or to Indigenous groups represented at the fair. This led me to investigate each occurrence on a case-by-case basis before assigning each mention geographical coordinates of a particular place in the world. Other cases where terms like “America” could refer to Latin America rather than the United States were also closely examined before the geocoding process. Further, close reading of a significant sample of the newspaper articles was important to understand particularities of the historical context of colonialism in some cases. For example, a few occurrences of “German” referred to “German East Africa,” the region of present-day Burundi, Rwanda, Tanzania, and part of Mozambique that, by 1904, was still under German colonial rule. Similarly, in the wake of the Russo-Japanese war that happened simultaneously to the fair, mentions to the region of “Manchuria” were often associated to the Japanese exhibit. Where it seemed fit, I have decided to assign general latitude and longitude coordinates of present-day Inner Mongolia – the true geographical target of imperialistic and expansionist policies of the Russian Empire and the Empire of Japan.

In the words of Shanon Leon, when it comes to humanistic inquiry, most data sets cannot “stand on their own without clear and thorough documentation that accounts for the many decision points along the way.”[[3]](#footnote-3) Further, Stéfan Sinclair and Geoffrey Rockwell have argued in favor of the interpretive responsibility of humanists and historians engaging with text analysis and quantitative methodologies. They remind us that computational tools do not produce meaning they are rather meant to “facilitate the augmented hermeneutic cycle,”[[4]](#footnote-4) and in this sense, without human intervention based on thorough knowledge of the input data and its historical context, the automated process of extracting named entities would have risked misrepresenting the experiences of particular groups that were objectified at the fair. With these and other ethical concerns in mind, the following provisions informed each step of the research process that resulted in *An Imagined Geography of Empire*:

* **Terminology:** Every effort should be made to avoid derogatory terminology in the deliverables of the project.
* **Awareness and Historical Accuracy:** Both in the narrative and in the rhetorical layering of data visualizations, the unavoidable presence of derogatory terms used to characterize Native Americans, Native Filipinos, and other Indigenous groups in 1904 should be explicitly marked as historical.
* **Misuse of Historical Evidence:** Since the core argument of *An Imagined Geography of Empire* is that local newspapers produced particular cultural representations of the world that centered the United States and subjugated Filipino culture, problematic language is expected to emerge out of the text mining process. The author acknowledges that some of the history unearthed in this project may be weaponized to perpetuate structural violence towards Filipino and Filipino-American people. *An Imagined Geography of Empire* is committed to communicating and debating this history ethically. The author invites descendants, Filipino-Americans, and other stakeholders of this history to get in touch with any concerns they might have in the future.
* **Sensitive Historical Material:** Where exhibition of historically sensitive photographs is deemed necessary for the argument at stake, there is an obligation on the part of the researcher to implement takedown procedures and remove any images in case any descendant, Filipino-American, or stakeholder reports the use as unfair, dangerous, misrepresentative, or violent towards them as individuals, or if anyone feels violated, unsafe, or prejudiced.
* **Acknowledge of Community and History Ownership:** The author acknowledges that, in October 2022, the city of Clayton, MO, initiated the process of officially recognizing and commemorating the site of the Philippine Village at the St. Louis World’s Fair as a site of misrepresentation, mistreatment, and violence perpetrated towards Native Filipinos in 1904. Since April 2021, Filipino-American artist Janna Añonuevo Langholz started to bring attention of neighbors and city officials to the matter and pressuring for recognition of the site and the commemoration of a new historical marker. In the last three years, the [Philippine Village Historical Site](https://www.philippinevillagehistoricalsite.com/) has been honoring and remembering the Indigenous peoples who were objectified at the fair. The author acknowledges the work that Langholz and other community members have undertaken for the preservation and communication of the Filipino memory and complicated history of violence and subjugation in St. Louis. *An Imagined Geography of Empire* further recognizes the Philippine Village Historical Site (PVHS) and its community as the leading force behind the production, representation, and sharing of knowledge about the Filipino American history, culture, and values. The author therefore is committed to aligning this project with the interests of the PVHS community. If any aspects of this work are considered for publication, every effort should be made to provide a representational body of the community with an opportunity to see, evaluate, and opine on the deliverables prior to official publication.
* **Issues of Copyright:** *An Imagined Geography of Empire* is committed to comply with Title 17 of the U.S. Code to ensure that photocopies are only used for the purpose of scholarship and research. Photocopies of The St. Louis Republic and the St. Louis Post-Dispatch were made available in Newspapers.com courtesy of Chronicling America and the State Historical Society of Missouri. Newspapers.com does not make explicit the original institution that provided photocopies of the St. Louis Globe-Democrat. Further, in accordance with the Library of Congress, *An Imagined Geography of Empire’s* use of these photocopies “fair use” and transformative as per Title 17 of the U.S. Code. Although newspapers published in the United States more than 95 years ago are in the public domain in their entirety, the JPG or PDF files retrieved from Newspapers.com will not be shared to the public as they can be found in both Chronicling America and Newspapers.com databases.

1. "Francis Honored Guest of World on Last Day of Exposition," *The St Louis Republic*, December 02, 1904. [↑](#footnote-ref-1)
2. “Igorrotes on Way to Island Home,” *The St. Louis Republic*, December 02, 1904. [↑](#footnote-ref-2)
3. Sharon Leon, “The Peril and Promise of Historians as Data Creators: Perspective, Structure, and the Problem of Representation,” *[Bracket] (blog)*, November 24, 2019. https://www.6floors.org/bracket/2019/11/24/the-peril-and-promise-of-historians-as-data-creators-perspective-structure-and-the-problem-of-representation/, 10-11. [↑](#footnote-ref-3)
4. Stéfan Sinclair and Geoffrey Rockwell, “Text Analysis and Visualization: Making Meaning Count,” in *A New Companion to Digital Humanities*, ed. Susan Schreibman and et. al. (John Wiley & Sons, Ltd, 2016), 345. [↑](#footnote-ref-4)