Tagging the Writer's Trace: How Computational Social Science Shaped (My) Freight Graffiti Research

Angel Abundis [1,*]
1 Independent Researcher, Guadalajara, México; E-Mail: abundiscomunicacion@gmail.com
* Corresponding author

Abstract
Graffiti on freight trains exists in a constant duality between the physical and digital realms, manifesting as works of art that circulate across the North American continent and are documented on user-generated content platforms. This phenomenon reflects the growing 'datafication' of everyday life and is explored through various theoretical perspectives and techniques.

Computational social sciences offer a valuable conceptual framework for analyzing this phenomenon. Instead of making up new concepts, researchers adapt existing frameworks to interpret data using computational tools and artificial intelligence, which allow for the aggregation and analysis of large volumes of data and the discovery of hidden patterns interpreted on solid conceptual frameworks. However, analyzing New York graffiti on freight trains requires specialized and niche approaches, involving data collection, custom model training, and the use of unconventional metrics.

This paper shares the author's experience as a researcher, showing how the perspective of computational social sciences guided an investigation into a type of graffiti and a geographically dispersed community. It discusses the theoretical and methodological decisions made, as well as the limits, scope, and possibilities of this approach.

Keywords
computational social science, freight graffiti, datafication, mixed methodology

1. Introduction

In this article, a series of reflections and personal experiences will be presented regarding the use of computational tools within the framework of social science research. This research resulted in the academic project "Visualización de un campo de interacción hipertextual en forma de red utilizando procesos computacionales. Caso de estudio: Graffiti en trenes de mercancía en norteamérica," with which I earned a master's degree in Communication from the University of Guadalajara.

This research culminated in a printed monograph and a couple of programs that, first, enable the backing up and, subsequently, the visualization of a hypertextual interaction field on Instagram. *Instagram Data Mining Bot* (idmb) was coded in Python; it is software that uses the open-source library *Instagrapi* (Andreev, 2020), which employs Instagram's Application Programming Interface (API) to interact with its database programmatically. Additionally, the Data Visualization Interface (DVI) is a control panel that enables the management, visualization, and interaction with the interaction field as a network graph (Figure 1), thanks to the *Graphology* (Plique, 2021) and *Sigma.js* libraries.

The technical, step-by-step explanation of how these network graphs are formed was presented in the talk "Node Reduction through Artificial Intelligence Inferences using Graphology and Sigma.js: A Case Study on Hypertextual Conversations in Freight Train Graffiti in the North American Region" at FOSDEM 2024, and in the article "Mining, Shaping, Visualizing, and Interpreting Instagram Hypertextual Networks of Freight Train Graffiti Communities in North America Using Custom Machine Learning Models and Graphology" (Abundis, 2023).

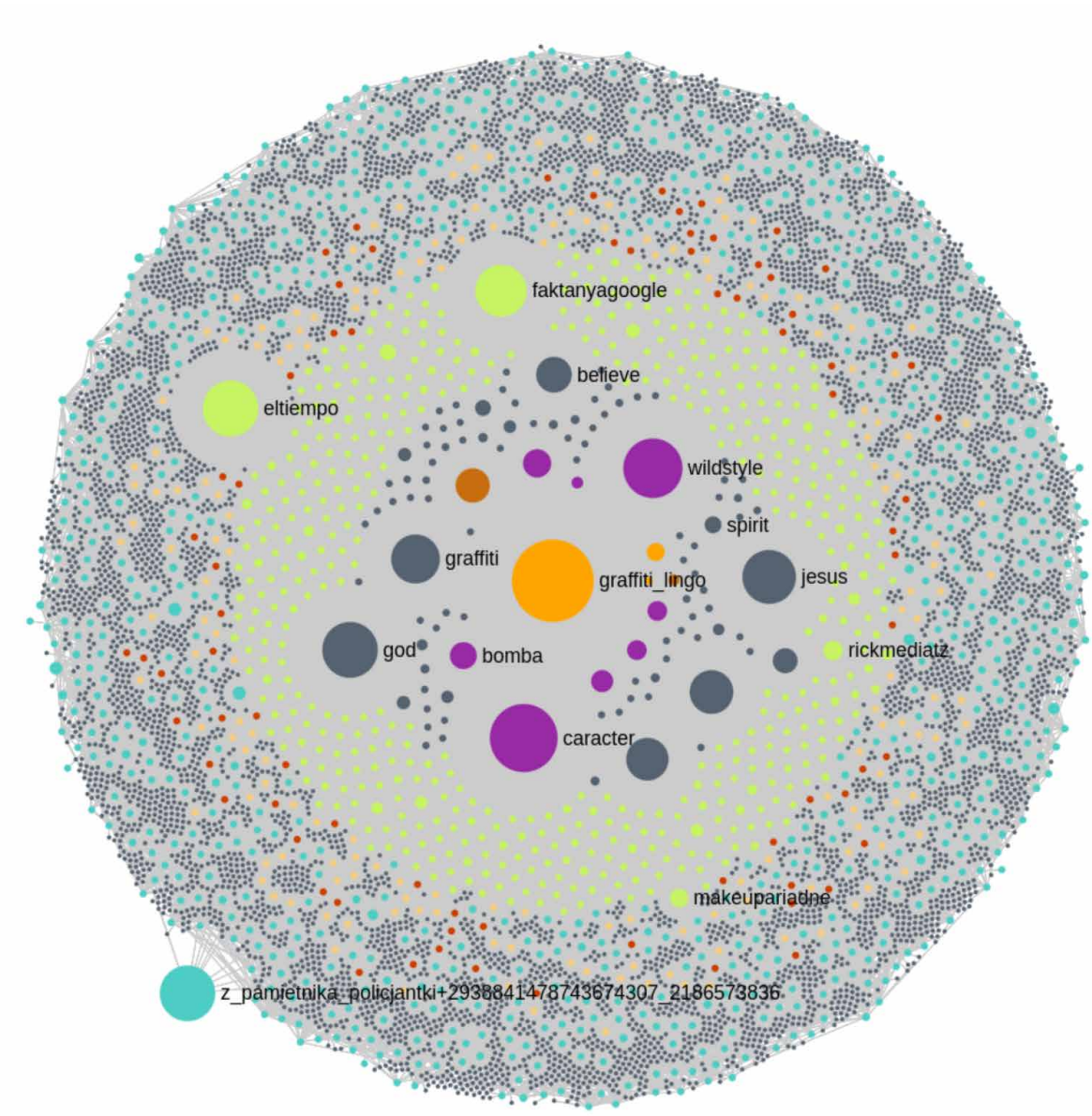

Figure 1. Network graph of the conversation: kosm_1_hashtagTop_9_62568b82. Source: Own elaboration, available at the following link: https://data.abundis.com.mx/vista/hashtags_ai_data_live.php?id=72

During the production of this dissertation, a key personal discovery was the adoption of computational social sciences as an analytical perspective. If summarized in a single phrase, it would be: the synthesis of a conceptual approach in social sciences with methodologies that leverage computational processes to their fullest potential. While this "maxim" may seem self-evident—as sociologists, anthropologists, and communication scholars increasingly integrate new technologies to optimize their workflows—the integration referenced here extends further. This approach aims to consider the influence of computational processes at each stage, as far as possible, within the research design, assessing their impact on conceptual frameworks, and, consequently, understanding the process in which a phenomenon is transformed into interpretable data.

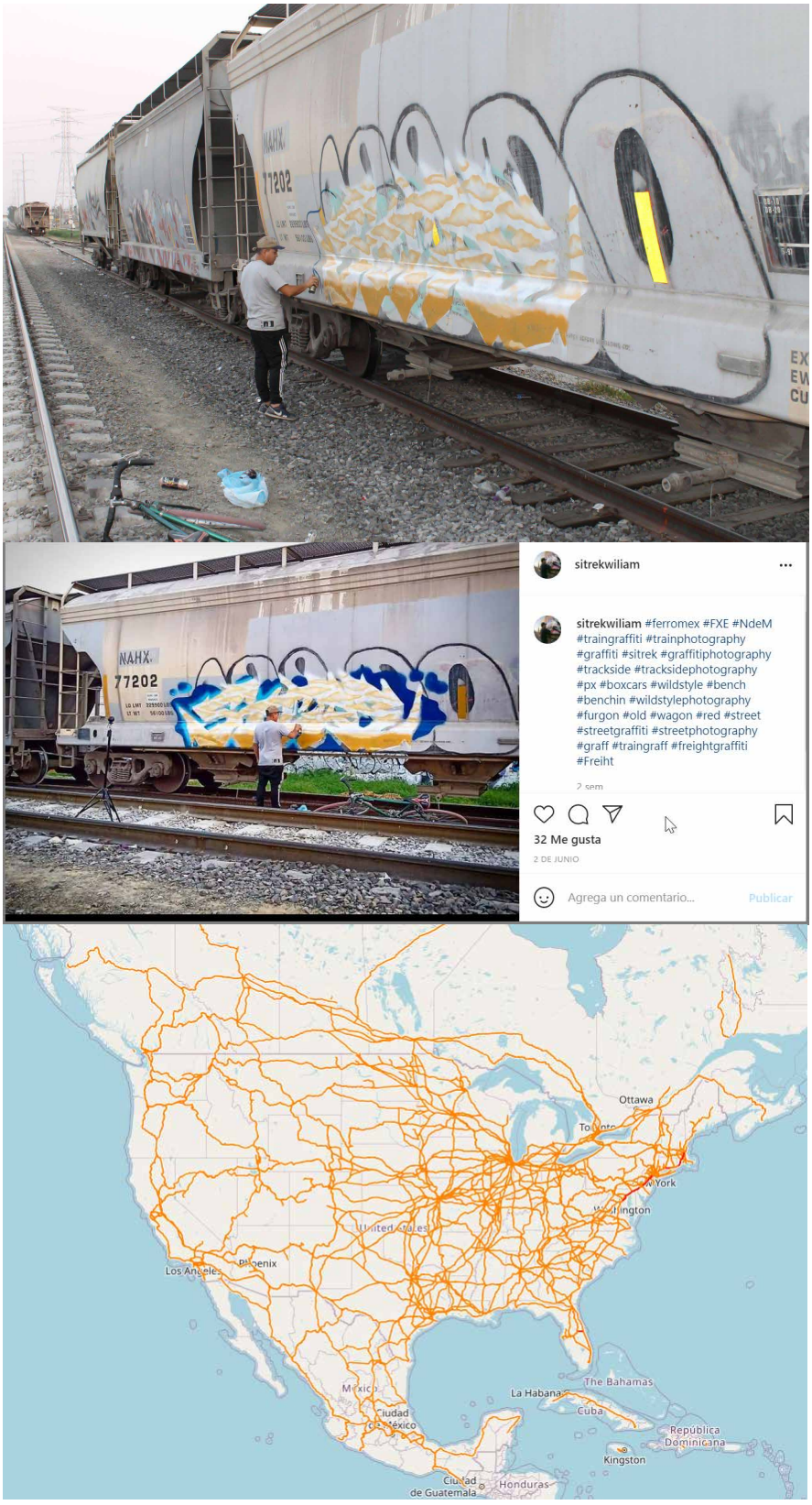

Figure 2. Circulation circuits of Citrek's intervention. Source: Own elaboration.

2. #Citrek #Jupiter #FreightGraffiti #Mexico #NorthAmerica

Freight graffiti uses rail vehicles as its canvas and the railway network as its circuit, in this case across the North American region, making it inherently mobile and transnational. A train tagged with a wildstyle piece by Citrek in the yard located in "San Juan" in Zapopan, Jalisco, Mexico, has the potential for transnational transit between Mexico, the United States, and Canada. Physical mobility allows freight graffiti to work as a symbolic bridge between local scenes, announcing the writer's presence to people from different regions, creating a transnational community of graffiti writers. This national symbolic circuit establish, after several years of local or regional transit, trains marked from both American coasts arrived in yards on the opposite one, in cities like New York, Los Angeles, and Chicago, which already had established graffiti scenes (Gastman et al., 2006, pp. 110-134).

NAFTA in 1994, a transnational trade agreement between Mexico, the United States, and Canada, unified domestic circuits. According to writer Jupiter, trains were already decorated in Mexico before the borders opened, but in 1999, after the privatization of the Mexican railway system was completed, transborder railway lines were consolidated. This allowed decorated American trains from cities north of the Rio Grande (whether American or Canadian) to circulate in Mexico and vice versa (Abundis, 2022, p. 204).

During one of the fieldwork excursions, Sitrek decorated a railcar with the economic number NAHX 77202. The rolling stock in question was parked, anchored alongside six other cars, on a railway track. Writers do not mark cars at random and recognize those that have international reach. Sitrek shares: "Hay vagones que son de aquí de México nomás, y hay otros que son los que viajan. Por lo regular, son los intermodales, los autoracks, tolvas gringas, las BNCF, Union Pacific (There are cars that are just from here in Mexico, and there are others that are the ones that travel. Usually, they are the intermodals, autoracks, American hoppers, BNCF, Union Pacific)" (Sitrek, personal communication, 2021). Once the work is finished, Sitrek smiles, satisfied with the wildstyle he has just completed, sending a clear message to the graffiti community: Sitrek is, exists, and was here.

A couple of days later, Sitrek posted on Instagram one of the photographs taken during the field visit. The post lacks any traditional caption or footer, meaning the image is not accompanied by any phrase or sentence. Instead, it consists of twenty-six hypertextual hashtags related to the graffiti world and concepts developed within this contextual framework. Railway companies such as Ferromex, acronyms FXE and NdeM (Nacionales de México), and words related to New York graffiti tradition like wildstyle, graffiti, street are featured. Terms such as boxcar, furgon, or trackside, which are part of the internal language of railway workers or rail enthusiasts, are also used. In addition to the writer's name and crew: Sitrek PX.

The use of these hashtags (hypertextual tags) allows the graffiti writer to publish their interventions and integrate them into broader conversations with other platform users who use any of the hypertextual tags. The writer selects hashtags to engage in broader conversations, joining circles where they can announce their participation to others and announce their presence in another communication channel within the group, akin to marking streets and freight trains. Therefore, the intensive use of hashtags, and the principle of connection that arises from this resource, enables linking both circuits (figure 2) in which Sitrek aims to gain visibility, or getting up (Castleman, 1984).

Finally, there are two groups that document and post interventions on freight trains. First, there are the writers who, after completing their works, photograph them and may or may not share them on their digital social networks. Second, there are the benchers, a broad term that actually encompasses two distinct groups: on one hand, rail enthusiasts, who are not directly linked to graffiti and focus primarily on trains. This group organizes meetings alongside the train tracks, records the train itself, and publishes them on digital video platforms, where writers seek their own interventions, which they then share on their personal profiles. Finally, there are freight graffiti benchers, who document interventions on trains in video or pictures and publish them on social digital platforms, tagging the writers (Abundis & Granizo, 2022, p. 126).

The bench is the central interaction that threads together the community on Instagram of New York graffiti writers and enthusiasts of train graffiti. Jupiter defines it as "la acción de tomar fotos o videos a los trenes, a los graffitis en los trenes (the action of taking photos or videos of trains, of graffiti on trains)" (Jupiter, personal communication, 2022). After documenting the pieces, he mentions, "la comunidad, por medio del hashtag o el arroba, te mandan las fotos. [...] A mí me ha mandado fotos de Rusia y de un chingo de otros lados del mundo. Me ha mandado fotos de Canadá, de EEUU, sobre todo (the community, through hashtags or mentions, sends you the photos. [...] They've sent me photos from Russia and many other places around the world. They've sent me photos from Canada, the US, mostly)" (Jupiter, personal communication, 2022). The freight graffiti benchers (hereafter referred simply as benchers) emerged as a key finding from this exploratory fieldwork. This valuable insight was gathered through semi-structured interviews, graffiti writer guided tours, and the recording of interventions with Jupiter and Sitrek from the PX crew. It is relevant because benchers play a pivotal role in the hypertextual interaction fields explored in this study, which presents a nuanced relationship with purely computational methods or rushed researchers, hence the balance between disciplines and fieldwork.

3. Computational Social Science

Lindgren's perspective (2020) offers a theoretical and methodological approach that seeks to integrate interpretive sociology with computational analysis. Lindgren prioritizes the use of classical theories such as those of Durkheim, Weber, and Marx. These are chosen for their conceptual complexity and established capacity to address social phenomena, contrasting with newer theories that may rename already conceptualized phenomena or conceptually isolate them from broader frameworks. Lindgren does not exclude contemporary conceptual frameworks from authors like Thompson, Latour, and Merton, emphasizing the importance of these frameworks being robust and interoperable, capable of adaptation and modification according to researcher needs.

Methodologically, Lindgren's perspective allows for the use of advanced tools in computational sciences. These include network analysis and visualization, digital data collection and mining, management of relational databases, and machine learning models inferences. This approach is particularly relevant in the study of contemporary phenomena where the use of social digital platforms and the datafication of society generate an unprecedented amount of data, given their inherent and vital role in contemporary daily life (Gomez Cruz, 2022). Lindgren (2020) refers to this when claim that the development of the Internet and social digital networks turns our everyday lives into data.

Adding the perspective of computational social sciences to this document is crucial because it moves us away from the idea of big data and artificial intelligence (AI) as black boxes into which a certain amount of data is fed and they always yield unquestionable results (Lindgren, 2020). Instead, it drives us to think on every methodological step, as each decision holds considerable importance. For instance, defining the scale of analysis, determining which types of metadata retrieve; or whether the dataset will consist of graffiti writers' posts, the list of profiles they follow, posts published under specific hashtags, tagged with certain locations, or all of the above (Abundis, 2023). In conclusion, the ability to position oneself from both fields indeed involves establishing a cross-field logic. This involves designing and executing code that not only processes data on a large scale but does so within a conceptual framework predefined by the researcher, taking into account both cultural context and symbolic patterns that often get lost in purely technical approaches that have not been finely tuned. Programming does not require mastery of computer science, but rather a commitment to adapt and understand computer science perspective from the social sciences point of view. This possibility, in my experience at least, was made possible thanks to the coding democratization brought by the consolidation and popularization of open-source philosophy.

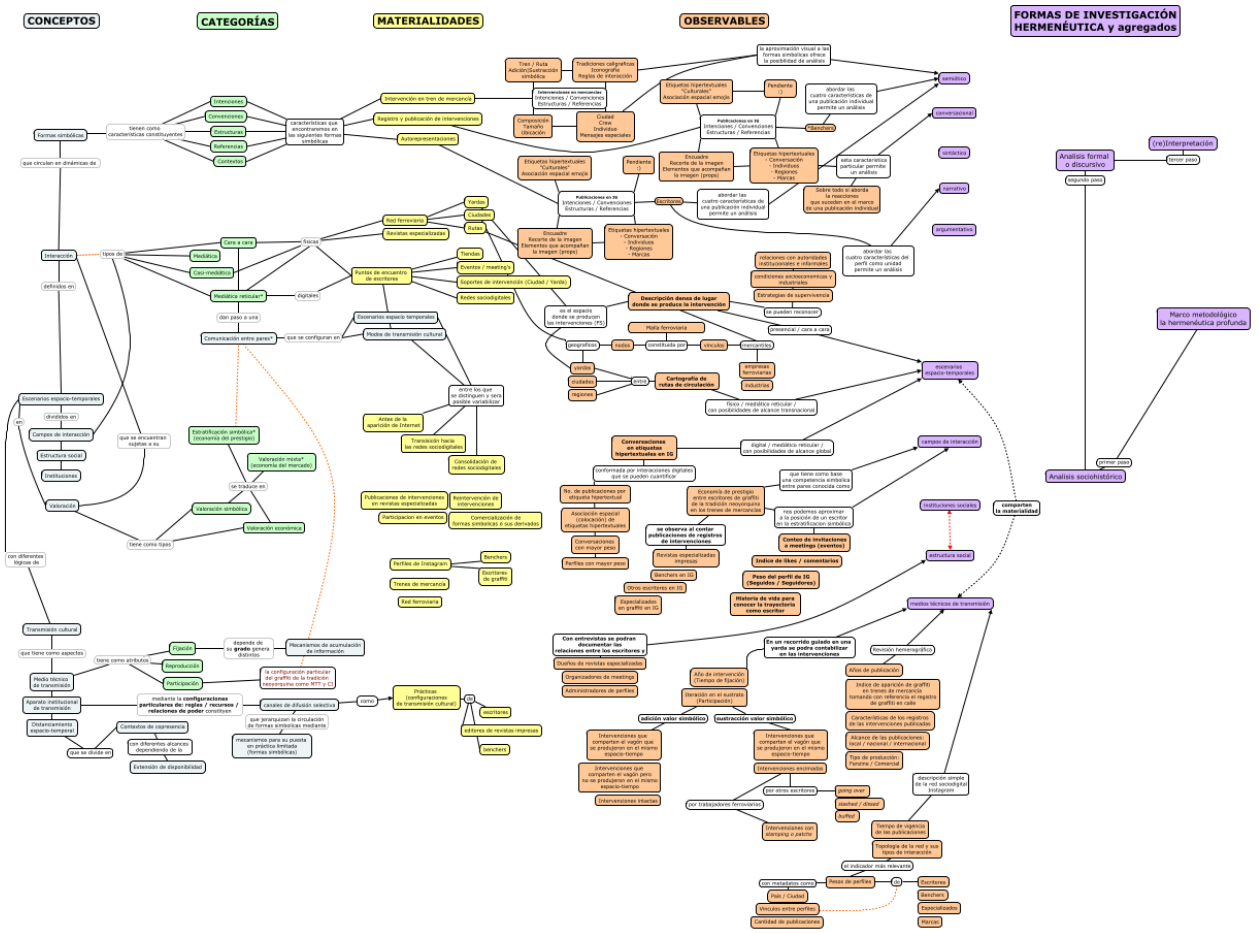

Figure 4. Concept map of Thompson's Mass Communication Theories (1991, 1998). Source: Own elaboration, available at the following link: https://data.abundis.com.mx/mapa-conceptual.pdf

4. Implications of Computational Processes on the Study Object

One of the most fascinating, yet challenging phase in this graduate program was shaping the object of study. The "translation" from social phenomenon to object of study involves various filters of refinement and synthesis across empirical, theoretical, and methodological dimensions.

The first step involved observing the phenomenon of graffiti on freight trains, which resulted in an exhaustive contextual framework in terms of quantity and diversity, aiming to explore all possible approaches to the social phenomenon. This exercise encompassed describing railway graffiti from various perspectives, particularly focusing on the Mexican context.I can assert that the decision to work with hypertextual conversations on Instagram was driven by the platform's content production principle, which enables the collection of meaningful data on a large scale. In Instagram users voluntarily share data that formed collaborative networks of meaning through the use of hashtags. Users actively engage in (and produce) symbolic content full of meaning, primarily through hypertextual tags. This symbolic content emerges and originates elsewhere: Writer names (entities) and their crews, the cities where train cars were marked or benched, graffiti glossaries, the type of rolling vehicule marked, and the rail company it belongs to are some elements that can be integrated into these tags and analyzed individually or as part of a collective conversation..

This leads us to the second step: defining the study object guided by a theoretical framework. Towards the conclusion of this work, I adopted J.B. Thompson's conceptual framework, following considerations of Bruno Latour, Benedict Anderson, and Carlos Scolari. This choice was driven by the belief that Thompson's theoretical framework facilitates the interoperability of concepts through the data it generates. A pivotal methodological exercise supporting this notion was the creation of a conceptual map within a methodology course (Figure 4). Divided into four columns organizing concepts, categories, materialities, and observables, a fifth column was added: "hermeneutical research forms," which categorizes the methods proposed by the English sociologist to organize his objectification and interpretation methods. In this exercise, it was confirmed that the Thompson's framework was strongly interconnected, considering that observables and materialities had many bridges between them, formed because they share categories of concepts and data manufactured from this conceptual key. In this case, the clearest example is the most simple one: the symbolic forms properties (intentions, conventions, structure, references, and context) are transversal and appear in different expressions, from the writer's intervention in a rolling vehicle to their publication on Instagram, and even in the retransmission of these graffiti in print magazines of the 1990s.

After selecting the conceptual framework, or possibly in parallel but as the third step in this text, comes the task of translating the social phenomenon into data. It is through methodological design that we outline the aspects of the social phenomenon that are of interest. The impact of computational social sciences at this stage lies in using various processes to fetch, mine, process, and ultimately model meaningful data for research. I really believe that techniques for data retrieval and storage significantly influence the selection of observables we choose to extract from the available materialities, alongside our choice of theoretical framework and concepts to work with. Furthermore, the technical and methodological possibilities depend on the researcher's ability to adapt computational processes to once again retrieve, store, process, and visualize the dataset.

5. Triad of self-announcement as an operational box

Thompson's (1991, 1995) mass communication theories align with Lindgren's perspective. They are a robust framework of interpretative sociology that allows for the interoperation of its concepts. Importantly, this framework is open in nature, as its modular design permits the recognition and modification of properties and attributes of these concepts. A theoretical overlap (see Figure 5) occurs between: 1) Thompson's symbolic forms (1991), defined as communicative and cultural expressions characterized by intentions, conventions, structures, and sociocultural references circulating within historically structured contexts; and 2) Figueroa's triad of self-announcement (2014), which asserts that signature grafiti (*grafiti de firma*) undergo sociohistorical reconfiguration through three vertices: I am (*yo-soy*), I exist (*Yo-existo*), and I was here (*Yo-estuve-aquí*).

The signature grafiti (*grafiti de firma*) as a symbolic form with intentions, conventions, and structurally organized references within specific socio-historical contexts becomes useful after linking the concept of the triad of self-announcement with systemic elements (Thompson, 1998, p. 211). A synthesis of this idea suggests that the creative impulse inherent in the three types of announcement can be observed, and turned into data, as systemic elements within symbolic forms. Considering graffiti in the New York tradition, it is argued that systemic elements appear in interventions by writers in the city, on passenger and freight trains, and their posts on user-generated content platforms, like Instagram. This indicates that the meaningful autoreference is cross-cutting and operates across all versions of graffiti as a symbolic form.

This triad is central, as graffiti writers use their name or crew as hypertextual references on sociodigital platforms; that is, the self-referential entity moves from the streets or freight trains (where it initially originated) to the sociodigital realm (the phenomenon addressed in this document). For instance, after documenting and publishing an intervened train, Jupiter tags his photograph with the hashtag #JupiterPX, announcing himself on another platform. As previously mentioned, the writer's entities (their name and crew) are not the only tags used in each post. We also find thematic hashtags, which refer to terms

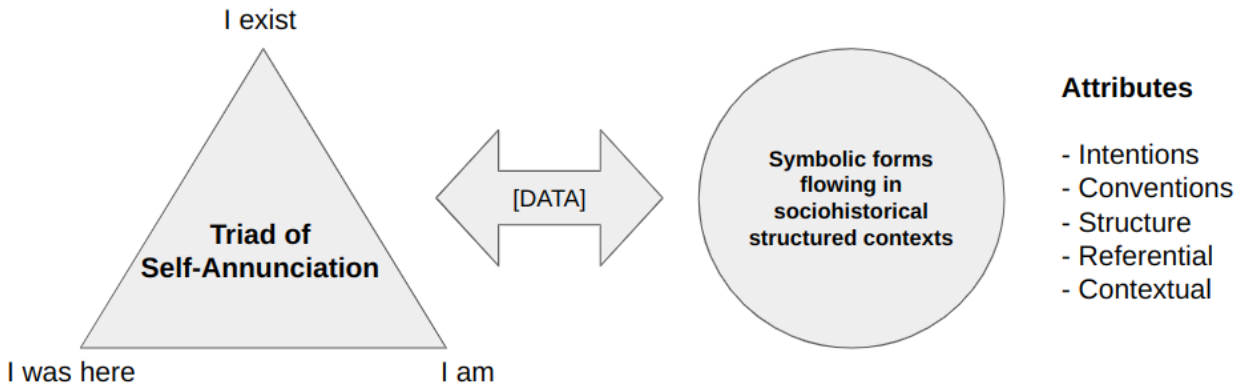

Figure 5. Overlap between Figueroa's triad of self-announcement and Thompson's symbolic forms.
Source: Own elaboration.

Figure 6. Figueroa's triad of self-announcement as an operational framework for categorizing hashtags.
Source: Own elaboration.

that make sense within the graffiti community on trains. Additionally, geographical tags allow the location where a rolling vehicle was marked or benched to be identified.

Using spaCy, two text analysis techniques were used to identify and categorize hashtags according to each vertex of the triad of self-announcement. I am, which identifies the names of writers and their crews, utilizes the Out of Vocabulary (Oov) technique, where the term is searched in a dictionary containing all known words in Spanish and English; if there is no match, the tag is then marked as a possible entity. To validate that it is an entity related to the phenomenon, the word "graffiti" is searched within the entire image description. Finally, depending on the number of characters, it is determined whether it is a writer's name or a crew. For the vertices I exist and I was here, the opposite logic is used. All known words within the hashtag are disassembled and searched in specific dictionaries for each vertex. If there is one or more matches, the relevant terms are marked and the type of dictionary it belongs to is indicated. For I exist, two dictionaries were used: a glossary of terms from New York graffiti and a dictionary of words used by railway workers. Together, these dictionaries form a collection of words that constitute the slang used by graffiti writers on trains. Finally, for I was here, a dictionary was compiled with all cities in the North American region with a population greater than 30,000 inhabitants. Figure 6 shows how this theoretical figure becomes an operational framework, using the logic outlined in this paragraph.

Recognizing entities, places, and relevant words is a straightforward task using the Named Entity Recognition (NER) component of the spaCy library, as long as traditional semantic documents are used. However, it fails completely when using image descriptions composed of hypertextual tags. By default, it does not remove the hashtag (#) from the tag, does not break down the words that make up a hashtag like #FreightGraffiti (freight, graffiti), and ultimately does not know how to handle independent terms, as it was trained to identify words in semantic positions. To solve this theoretical challenge, algorithms were developed to parse tags and identify these terms. While not a conventional machine learning model due to technical constraints, this approach, refined through multiple iterations and patches, ultimately achieved positive results.

6. There is no ready-made AI models for your study object

Based on past experience, it's acknowledged that existing Artificial Intelligence (AI) models do not directly solve researchers' specific concerns. Instead, these computational processes must be tailored, coded, or modeled to meet the researcher's particular needs, always within the constraints of their technical and financial resources.

Graffiti type (tags, throw-ups, rollers, wildstyle, characters, monikers) on freight trains is a crucial indicator in gauging the prevalence and variety of graffiti discussed in hypertextual conversations. This study employed a really simple machine learning model trained to identify these graffiti styles in posts previously collected through *idmb*. Each graffiti style was represented by 150 images sourced

Figure 7. Same image processed with different object detection models, the first trained to identify everyday objects, and the second trained with types of graffiti. Source: Own elaboration.

from relevant hypertextual tags such as #tagsandthrows, #wildstylegraffiti, #3DGraffiti. Images not fitting the graphic style or meeting minimal lighting conditions for model training were filtered out. Using the LabelImg tool (Tzuta Lin, 2015), each graffiti piece was labeled with a specific graffiti type boxbound tag. The resulting dataset comprised 664 images, used to train an object detection model with TensorFlow. Figure 7 illustrates the outcomes when processing the same image with different object detection models: on the left, a Google model trained for real-world object recognition; on the right, the model specialized in identifying graffiti types. While reliable within its specific scope, this model could potentially be enhanced by integrating a suite of models that collaborate to identify graffiti types, colors used, graphic elements, stylistic lineages, and other relevant characteristics.

I believe that research from a pure social sciences perspective faces both a challenge and an opportunity due to the lack of pre-designed artificial intelligence models for specific social phenomena, such as graffiti on freight trains. Creating specialized datasets and models is a collaborative endeavor that necessitates collective effort among researchers. This approach is expected not only to enhance the quality and quantity of data but also to improve model accuracy by incorporating diverse perspectives and experiences. Moreover, adopting open licenses for them should be a fundamental principle. This openness facilitates access and reproducibility, enabling other researchers to adapt and refine the models, thus contributing to progress in each field.

6. Research results as proofs of concept

Up to now, I've discussed the impact of integrating computational social sciences as a perspective in research. While I focused on constructing the study object and its objective tools, this segment will now explore how computational social science offers unique possibilities to share and display our research.

The most significant opportunity lies in the ability to share technical processes involved in data acquisition, compilation, analysis, and visualization as practical, executable exercises. While understanding these computational steps may require some basic programming logic, sharing each process as an executable allows for enriching discussions on specific social issues and their research. Proof of concepts is integral to social sciences, forming the core of disciplines seeking peer-validated knowledge. However, there are now alternative formats diverging from traditional printed monographs, enabling researchers to document methodological foundations, fostering debate and engaging other interested parties.

I believe that traditional printed text as the final delivery format imposes technological limitations on research that extensively employs computational tools throughout the process. This can potentially constrain the validation of theoretical and methodological frameworks and delay their evolution. There are various open-source tools that allows executable proofs of concept in computational social sciences directly from the browser. The most relevant is Jupyter Notebooks, which enable the installation of libraries, execution of heavy processes in interpreted languages like Python or R (including artificial intelligence libraries like TensorFlow), visualization of results using libraries like Matplotlib, and let the researcher comment out each step. Another perspective focuses more on data visualization rather than modeling: platforms such as Tableau or Power BI, where databases, in varying stages of processing, are ingested to create visualizations that validate researchers' arguments. Lastly, the third option, utilized in this academic work, involves using JavaScript libraries to custom code visualizations, algorithms and processes.

To illustrate the integration of browser-executable technical processes in supporting a hypothesis or assumption, I will focus on Jupyter Notebooks. According to its official documentation, a computational notebook serves as a collaborative document that integrates computer code, plain language explanations, data, and advanced visualizations such as 3D models, tables, charts, and interactive controls. These notebooks enable the addition and execution of code to observe real-time outcomes. As a result, digital scholarly works can include practical demonstrations at every research stage, starting from observing social phenomena and designing studies with small data samples, to validating theoretical assumptions.

During the master's program, there were at least two assignments with these characteristics. The first one was a hybrid between a monograph and a controversy map, which from Latour's perspective aimed to list all the actants involved in the yard at San Juan de Ocotán, where Citrek tagged a freight car. These concerns already appeared in this school assignment, as one of the paragraphs argued: "This homework is not a fully formed controversy site, nor is it a formal monograph; it is a hybrid, a chimera of transition. It does not provide a robust platform to be considered a site with hypertextual categories for nonlinear navigation. Instead, it consists of a very localized observation focused on territory and a specific moment that does not follow the basics of a site of the magnitude proposed by Latour for a controversy site. Rather, it adopts the linear format of a written monograph with added possibilities for the reader: data without interpretation in the interface, the interpretation of tools in use, and most importantly, the option for interaction and reinterpretation." [Translated from spanish]

In this work, titled "Descripción de sujeto a investigar: Escritor de graffiti de trenes de carga desde una interpretación libre de la teoría-actor-red de Bruno Latour (Description of the Subject of Investigation: Freight Train Graffiti Writer from a Free Interpretation of Bruno Latour's Actor-Network Theory)", and continuing in this idea, several multimedia resources were used to document an ethnography that could be accessed online and was interactive (see Figure 8). Most of the participant observation was recorded, including Citrek's intervention on the freight car and the yard in general, using an omnidirectional camera, providing readers with an immersive experience in the yard. As part of the mapping process, the open-source uMap service was used to identify and locate the actors featured in the fieldwork: The places, events, and people mentioned by the writers to the industries, garages where rolling stock could be intervened, or relevant places for the research.

Figure 8. Screenshots of interactive school work. Source: Own elaboration.

The second tool is DVI, the software developed to visualize hypertextual interaction fields, which allowed us to identify at least three dynamics within the context of freight train graffiti: divergence, convergence, and emergence. Divergence occurs when discussions veer into unexpected topics, sometimes due to the use of polysemic tags and other times due to generic tags. On the other hand, convergence is crucial as it shows users gathering around shared symbolic elements, consolidating community through common practices such as graffiti types identified by the object detection model or communal tags. Lastly, emergence underscores the need to be attentive to new significant terms emerging in the dataset, which might be overlooked in analysis as they do not belong to dictionaries made arbitrarily by the researcher for text analysis.

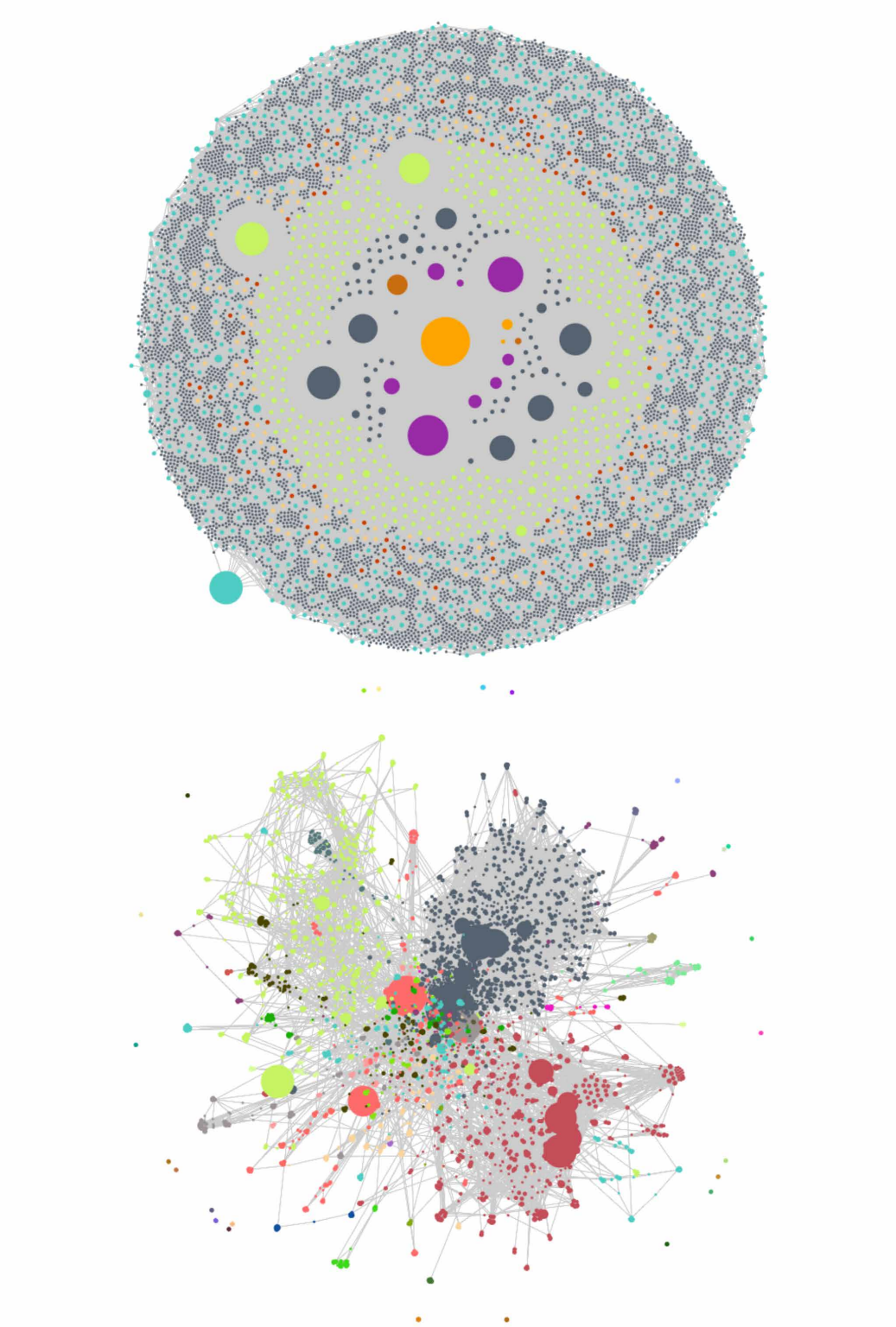

Figure 9. Visualization of the hypertextual conversation kosm_1_hashtagTop_9_62568b82. On the left, default visualization; on the right, after applying the ForceAtlas2 and Louvain Communities algorithms. Source: Own elaboration, available at the following link: https://data.abundis.com.mx/vista/hashtags_ai_data_live.php?id=72

These dynamics are present in hypertextual conversations, but they become evident when interacting with the data. A clear example is the polysemy of the term “Kosm”; the default visualization is a circle where some nodes are more relevant than others. However, it is only after applying the forceAtlas2 spatial layout algorithm and the Louvain community detection algorithm to color the groups that it becomes evident that there are three major thematic clusters.

In the network graph on the right (Figure 9), nodes in blue represent the subset of users and posts from the freight train graffiti writers’ community. This includes hashtags with names of Mexican freight graffiti writers like Siber, Aser7, Neika, and Kosm, as well as American writers such as Mecro, Jaber, and Ichabod. It also features communal hypertextual tags like #freightgraffiti, #traingraff, or #benching, which have been identified and marked. In burgundy, located in the lower right of the graph, posts from the user @kosm_world have been grouped. This user has formed a significant subset with his posts and the tags used, such as #jesus, #god, #power, and #faith. This user has given a different meaning to the term “kosm,” accompanied by a second community of users with similar interests who use a range of hashtags to share spiritual values. Finally, in lime green, another set of posts is grouped where the Polish term “kosm” (abbreviation for kosmetyki, cosmetics) is used in conjunction with tags like #crueltyfree or #weganskiekosmetyki. Reviewing this subset suggests it belongs to vegan activists promoting vegan cosmetics as an alternative lifestyle choice. Such findings are only possible through data interaction and can be shared using screenshots, like those in Figure 9, and applied in written form, but clearly, the possibilities are limited by the format.

6. Conclusion

In conclusion, this research reflects a journey that blends computational social sciences with interpretive sociology to analyze freight graffiti—a complex social phenomenon spanning digital and physical realms. By employing a conceptual framework informed by Thompson’s mass communication theories and Figueroa’s triad of self-announcement, the study navigates graffiti’s symbolic expressions, from its origins in physical spaces to its perpetuation through digital networks.

I find that combining computational social science with traditional sociological inquiry has offered a distinctive lens through which to understand the complex world of freight graffiti. Freight graffiti, with its roots in physical tagging on North American trains, expands its influence and impact as it travels through digital spaces. This movement across borders and platforms has transformed graffiti into a hypertextual form of interaction that connects local scenes and builds a transnational community.

My experience with writers like Jupiter, who tags internationally traveling train cars, provided a clear example of this dynamic. Jupiter’s use of Instagram to share his work, alongside hashtags like #JupiterPX and #freightgraffiti, serves as a form of self-announcement, where graffiti writers gain visibility and interact within a global community. It’s fascinating to observe how, by marking both physical and digital spaces, these artists expand their reach, leaving traces that blend art and identity in ways that defy borders.

Using computational tools to examine this phenomenon allowed me to address both the symbolic and practical elements of graffiti culture. By developing Python-based software and using machine learning models, I was able to classify various graffiti styles—tags, wildstyle, and 3D graffiti—capturing the diversity of content shared online. Data visualization tools provided a way to observe broader patterns in hashtag use, revealing how digital platforms serve as a cross-regional narrative for freight graffiti, where individual stories and styles converge into a collective digital tapestry.

This research journey has reinforced the power of computational social sciences to extend beyond purely technical insights, allowing me to engage deeply with the cultural and symbolic nuances of graffiti. As result of blending open-source technology with traditional social science approaches.

Conflict of Interests and ethics

The author(s) declare no conflict of interests. The author(s) also declare full adherence to all journal research ethics policies, namely involving the participation of human subjects anonymity and/ or consent to publish.

Acknowledgements

Dr. Figueroa's personal insights and scholarly work have highlighted the importance of viewing graffiti not just as a social phenomenon, but as a meaningful human expression. His concept of self-announcement has profoundly shaped my project, offering a new perspective on the topic. Beyond his academic contributions, I deeply value his supportive guidance and precise feedback. The message inscribed in *Graphitffragen*, a gift from him, resonates with me: Llos senderos de la historia reúnen a quienes se preguntan el mañana y recuerdan el ayer", and I really hope so.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., & Zheng, X. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*.

Abundis, A. (2022). Y sin embargo se mueve: Aproximación al concepto de graffiti en tránsito desde el graffiti en trenes de mercancía en la ciudad de Guadalajara. *Ñawi*, 6(2), 193–212. https://doi.org/10.37785/nw.v6n2.a11.

Abundis, A. (2023). Mining, Shaping, Visualizing, and Interpreting Instagram Hipertextual Networks of Freight Train Graffiti Communalities in North America Using Machine Learning Custom Models and Graphology. *UXUC - User Experience and Urban Creativity*, 5(2), 68-87. https://doi.org/10.48619/uxuc.v5i2.838

Abundis, A., & Lopez, C. (2022). El freight graffiti, un circuito de circulación simbólica transnacional oculto en el margen de las vías del tren. In Cuevas, A., Arellano, A. C., & Cortés, E. (Eds.), *Jóvenes, significaciones y escenarios de riesgo: experiencias y narrativas diversas*, 117-134. Universidad Autónoma de Tlaxcala.

Andreev, M. (2020). *Instagrapi - Unofficial Instagram API for Python*. GitHub. https://github.com/subzeroid/instagrapi.

Castleman, C. (1982). *Getting Up: Subway Graffiti in New York*. The MIT Press.

Gastman, R. (2006). *Freight Train Graffiti*. Harry Abrams.

Gomez Cruz, E. (2022). *Tecnologías vitales: pensar las culturas digitales desde Latinoamérica*. Universidad Panamericana.

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). *spaCy: Industrial-strength Natural Language Processing in Python*. https://doi.org/10.5281/zenodo.1212303.

Jacomy, M., Bastian, M., & Heymann, S. (2009). Gephi: An open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

Lindgren, S. (2020). *Data Theory: Interpretive Sociology and Computational Methods*. John Wiley & Sons.

Plique, G. (2021). *Graphology, a robust and multipurpose Graph object for JavaScript*. Zenodo. https://doi.org/10.5281/zenodo.5681257

Thompson, J. B. (1991). *Mass communication and modernity: Essays in critical theory*. Oxford, UK: Oxford University Press.

Thompson, J. B. (1995). T*he media and modernity: A social theory of the media*. Cambridge, UK: Polity Press.

Thompson, J. B. (1998). *Ideología y cultura moderna*. UAM-X.

Tzuta Lin. (2015). *LabelImg*. Github. https://github.com/tzutalin/labelImg