Data Visualizations as Propaganda: Tracing Lineages, Provenance, and Political Framings in Online Anti-Immigrant Discourse

NINA LUTZ*, PRIYA DHAWKA*, and KATE STARBIRD, University of Washington, USA

Along with other visual content, data visualizations are increasingly used within online discourse, including political communication. Though often considered to be "objective", data visualizations can also be created and/or appropriated to mislead. Here, we study the use and evolution of data visualizations within social media discourse around the ongoing "crisis" at the US-Mexico border in 2024. Through computationally-assisted qualitative analysis, we first describe how data visualizations are used to support four anti-immigrant frames, highlighting key tactics and sources of these visualizations. Next, we conduct a deep analysis of three Data Visualization Lineages (DVLs), exploring the role of adaptations, annotations, and remixing within families of data visualizations that share the same origin but have diverged through distinct visual alterations. We conclude by discussing approaches for supporting researchers in identifying and unpacking data visualization lineages, and highlighting design opportunities for mitigating the impact of misleading data visualizations in online discourse.

CCS Concepts: • **Human-centered computing** → **Social content sharing**.

Additional Key Words and Phrases: Social and Crowd Computing, Methodologies and Tools, Data Visualizations, Sensemaking, Crisis Informatics

ACM Reference Format:
Nina Lutz, Priya Dhawka, and Kate Starbird. 2025. Data Visualizations as Propaganda: Tracing Lineages, Provenance, and Political Framings in Online Anti-Immigrant Discourse. In *Proceedings of 28th ACM SIGCHI Conference on Computer-Supported Cooperative Work & Social Computing (CSCW) (CSCW '25).* ACM, New York, NY, USA, 47 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Disclaimer: *This paper contains data visualizations used to disparage migrants and asylum seekers in the United States. This paper also contains images of racialized language and mentions of violent crimes (both real and unsubstantiated). This content may be distressing to some readers.*

On July 13, 2024 at a campaign rally in Butler, Pennsylvania, a shooter fired eight shots at Donald Trump as he spoke from the stage, wounding the former President along with several audience members, and killing one [124, 136]. Some alleged that the tragic shooting had been motivated by harsh rhetoric from Democrats during a heated Presidential campaign [19, 80, 103]. Though the evidence for that is limited, there is no doubt that the hotly contested election included troubling

*Both authors contributed equally to this research.

Authors' Contact Information: Nina Lutz; Priya Dhawka; Kate Starbird, University of Washington, Seattle, Washington, USA.

CSCW '25, Oct 18–22, 2025, Bergen, Norway

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

rhetoric. As the shots were fired, Trump pointed to a large display featuring a misleading infographic about U.S. immigration [44, 73]. That graphic was likely familiar to many in his audience, as it had been spreading — and evolving — on social media for several months prior. This paper tells the story of that chart (and others).

Fig. 1. Donald Trump speaking at a rally in Butler, PA with a chart about immigration moments before an assassination attempt was made on his life [45]. Photo by Evan Vucci, Associated Press.

As the internet becomes increasingly visual and data-driven, data visualizations such as Trump's immigration chart become more prominent. With widely available data and easy-to-use tools for generating and marking up graphics, the creation of data visualizations — including ones that can be used to support misleading claims and propaganda that scapegoat, demonize, and/or dehumanize marginalized populations — is democratized, similar to news production more generally. Social media platforms render the production of these graphics participatory, as other users provide feedback and remix graphics as they spread. Data visualizations are often assumed to be objective and positioned as "factual evidence" in public discourse [113]. We explore how this "evidence" can be created and shaped to fit particular political frames, including ones that mislead and dehumanize a vulnerable population, in this case immigrants.

We apply computationally-assisted qualitative methods, leveraging computer vision techniques and classification in tandem with deep qualitative analysis, to study the production, evolution, and spread of data visualizations related to migration on social media. Throughout 2024, the U.S. was dealing with a so-called "border crisis" that included record numbers of asylum seekers, strains on the systems and communities that supported new immigrants, and criticism of U.S. immigration policy. 2024 was also an election year, and as it progressed, Republicans and their party's eventual nominee, Donald Trump, increasingly leveraged anti-immigration sentiment to criticize the current administration, scapegoat minority groups, and garner votes. Our work investigates visualizations that were part of deeply politicized rhetoric that often employed misleading content and propaganda in service of political advantage.

We ask two related research questions:

- **RQ1:** How do creators and social media users frame data visualizations on social media in political conversations?
- **RQ2:** How do content creators, politicians, and social media users generate and adapt these visualizations in ways that support, extend, or otherwise change these frames?

Through a mixed methods analysis of data visualizations from the social media site X (formerly Twitter), we investigate how visualizations support various rhetorical and political frames in online discourse about the US-Mexico border crisis. We introduce the concept of **"data visualization lineage"** to describe how social media creators, users, and political figures employ a variety of strategies to modify and blur the provenance of data visualization artifacts. We illustrate how these visualizations and particular modifications of them are used as "evidence" to support, and occasionally counter, anti-immigrant political frames. Finally, we contribute a set of forward-looking research opportunities and design recommendations in this space.

2 Background of the Border Crisis

In 2014, President Obama declared rising immigration at the US-Mexico border a humanitarian crisis [142]. This crisis persists today, with factors like climate change and sociopolitical unrest leading to unprecedented global human displacement and migration [48, 50, 96, 157]. This migration, primarily from non-WEIRD (Western, Educated, Industrialized, Rich and Democratic) countries, is expected to increase over time [100, 144].

In May 2023, the Biden-Harris administration ended Title 42 [185], a COVID-19 pandemic measure that allowed the US to turn away asylum seekers at the US-Mexico border. Following the lifting of these restrictions, migration at the Southern border surged and prompted media and public figures (including politicians) to label it "Biden's Border Crisis" [174].

People respond to crisis events by converging — increasingly online — to seek and share information [97, 149], part of a collective effort to make sense of high stakes information in an uncertain and dynamic information landscape [114, 199]. In the case of the 2024 US border crisis, ambiguity around immigration reform and its ability to solve the very real and complex human migration crisis, anxiety about an overburdened social safety net [14, 59, 146], and a heated election created conditions that were ripe for anti-immigrant rumor and propaganda. This nationalist discourse in the US resonated with similar conversations in Europe, where anti-immigrant sentiment due to an influx of migration has also been rising [18, 77, 86, 138]. This troubling rhetoric has had real-world consequences, including riots [99], shootings [15], and targeted violence and harassment against migrants [85, 99, 151] and those who provide support for migrants in the US [31, 101, 104]. Scholars have observed how online anti-immigrant rhetoric is connected to instances of this violence [85, 99] and bears negative health outcomes for immigrant communities [133].

Throughout 2024, several events sparked subsequent media coverage and online conversations. In January, an incident in Eagle Pass, Texas caused a dispute between the Department of Justice, Texas National Guard, and border patrols of the US and Mexico [58, 63]. In March, GOP Senator Ron Johnson and colleagues held impeachment hearings for Homeland Security Secretary Mayorkas over increased migration at the US-Mexico border [8]. In April, several reports about increases in Chinese migration fueled online speculation and discourse [11, 140]. Throughout the summer, criticism of Democratic Presidential nominee Kamala Harris's role as "border czar" [188] proliferated online. In September, Donald Trump repeated viral (and false) claims about Haitian migrants eating pets [105] during a Presidential debate. We focus on online discourse around events like these about the US-Mexico border as a backdrop for exploring the role that data visualizations play in this high-stakes and politically-charged discourse.

This paper was written in October 2024, but we acknowledge the ongoing and ever-evolving nature of the immigration crisis in the US and other nations.

3 Related Work

We summarize relevant research related to sensemaking, visual media as evidence and within problematic information, and data visualization.

3.1 Sensemaking as Evidence and Frames

We conceptualize the US-Mexico border discourse as a crisis and build upon prior CSCW work [36, 175, 176] in crisis informatics, a field pioneered by Palen and colleagues, which draws on multiple disciplines to understand how people use information in times of crises and uncertainty, often online [148–150, 158]. Crisis informatics has been extended to visual media [36, 123] and in larger and ongoing crises such as climate change, or in our case, mass migration sometimes exacerbated by many sociopolitical and environmental factors [96].

One foundational property of crisis informatics is the collaborative work of sensemaking [123, 190]. Sensemaking occurs when people come together to make sense of available information under conditions of uncertainty and ambiguity [182]. Several researchers have approached the study of online rumoring as collective sensemaking [61, 107, 128, 160, 178, 199]. Starbird and colleagues recently applied the 2007 data-frame theory, by Klein and colleagues, that explains how sensemaking takes shape through interactions between available data (or "evidence") and frames [110] to the study of online rumoring in crisis events [177, 181]—which we also adopt.

Framing, introduced by sociologist Erving Goffman, explains how people make sense of experiences by organizing them into existing mental frameworks [91]. This foundation was iterated upon by scholars like Robert Entman, who clarified the role of frames in information and media systems, particularly how frames define problems, diagnose causes, make judgments, and suggest solutions toward a unifying theory of the evidence in the frame [81]. In 1999, Scheufele extended this to how media frames shape and are shaped by public opinion [169]. Benford and Snow followed with how framing is utilized in social movements and mobilization [29, 30]. These various theoretical frameworks have been actively applied to several social media studies [84, 112, 129, 155, 159, 183, 195, 198], including visual media.

In this work, we seek to apply the lens of collaborative sensemaking and participatory media from Starbird et al [177–179, 181] and other crisis informatics scholars [184, 193, 199] to the study of visual media — in particular data visualizations of the border crisis. In doing so, we draw upon Entman's framework [81], due to its applicability to political discourse, to analyze the collaborative "work" of producing and remixing content [178, 179, 193] or "evidence" to fit prevailing frames within immigration discourse.

3.2 Visual Media as Evidence

Media scholars have long studied how images have historically had and maintained a particularly powerful role as a centralized authority of truth and proof [27, 65, 75]. Visual evidence, and its impacts, are rooted in human psychology, with researchers observing that claims supported by images are deemed more trustworthy by participants [170]. Scholars have also investigated the role of imagery in misinformation campaigns [130] and the role of visual media in propaganda, particularly in social media [26, 62, 115, 145, 171, 172]. Concerns of these impacts have intensified with generative AI [93, 152], which can create images quickly and in large quantities and target political discourse, as seen in election disinformation campaigns in India [186].

The stakes of visual media have driven efforts to trace image provenance through media forensics, analyzing both image pixels [34, 82, 83, 132, 189] and metadata [35] to track creation and manipulation. Though social media platforms have yet to adopt provenance standards, some companies are implementing Coalition for Content Provenance and Authenticity (C2PA) guidelines for synthetic media [1]. Recent work explores how provenance affects user trust and interaction with online media and how technologists may build better tools and standards for doing so [84].

Interest in understanding an image's creation, and what ideologies may be embedded in this creation, are not new [27]. Images today are not just created by professional cameras but by ubiquitous smartphone cameras, computer programs (Adobe Photoshop, Microsoft PowerPoint), and artificial intelligence models (such as GANs, text to image generators), which scholars have observed are impacted by positionalities and choices of their designers, datasets, and even visualizations of data itself [37, 153, 164, 167, 194]. In this work, we examine these visualizations as pieces of visual evidence that undergo participatory dynamics, and changes, as elicited in §3.1.

3.3 Data Visualization

Previous work in visualization research has focused extensively on topics such as graphical perception and perceptual accuracy (how accurately individuals perceive visual representations of data) [49, 191], cognitive functions and biases [70], and design conventions to maintain objectivity [106, 113]. Due to this conventional association with objectivity, data visualizations are used in various fields to make evidentiary claims, communicate objective facts, and support analytic decision-making, among many other functions [68].

However, the visualization research community is increasingly questioning the notions of objectivity associated with data visualizations. Dörk et al. [76] argue that designing visualizations is an inherently subjective and political process, advocating for designers to take a critical approach that "promotes disclosure, plurality, contingency, and empowerment". Similarly, Correll [60] highlights the ethical responsibilities of visualization researchers and designers. D'Ignazio and Klein [69] consider data (and by extension the resulting visualizations) as being situated and subjective, and potentially influenced by the individuals creating, collecting, and working with the data. Lee-Robbins and Aydar [119] emphasize that designing data visualizations is a value-laden process by broadly outlining the motives to communicate affect to viewers. In response to positivist visualization models that emphasize the objectivity of data visualizations, Berret and Munzner recently proposed an interprevitist model of sensemaking for visualization [33]. These various debates about the objectivity of data visualizations often do not sufficiently consider how the assumed objectivity of data visualizations are currently being (re)-appropriated by malicious actors in the spread of online misinformation.

Visualizations have great potential to mislead audiences through specific techniques for lying with charts [47] and have a long history of being a part of disinformation campaigns, as Bergstrom and West note [32]. As more tools to create data visualizations emerge and their presence in problematic information campaigns become more spread, the visualization community is increasingly interested in understanding how online communities create, share, and use data visualizations and how these artifacts contribute to misinformation campaigns. Lee et al. [116] examined how coronavirus skeptics engage with conventional visualization techniques to create what the authors describe as "counter-visualizations" that challenged mainstream narratives around data and spread misinformation during the COVID-19 pandemic. Additionally, Holder and Xiong-Bearfield [94] explored the potential of data visualizations for reinforcing polarizing political attitudes among Republican and Democratic viewers.

Studying the role of data visualizations in misinformation campaigns, Lisnic et al. [122] introduce the concept of "vulnerable" visualizations, which, through inadequate context and biased framings, are misappropriated by online audiences to spread misinformation. Furthermore, Lisnic et al. [121] conducted a mixed-method analysis of tweet responses to misleading and vulnerable visualizations and found that social media users often point out the inaccuracies and fallacies in these kinds of data visualizations. Although Lisnic et al. [122] propose a number of research opportunities around data-driven misinformation, their work focuses on a quantitative understanding of how data visualizations contribute to misinformation. In this paper, we contribute a qualitative analysis of how social media creators iterate on and co-opt open data sources and data-driven visualizations to spread anti-immigrant sentiment in online political conversations.

4 Methods

We employ a grounded, interpretivist approach [53, 54] — supported in key places by computational tools — to sample and analyze visual media (in this case, data visualizations) shared on X related to the US-Mexico border crisis. This approach adapts similar methods employed in studies of online rumors [127, 182] and political discourse [130, 159, 180] to the study of visualizations shared through social media [116, 121, 122].

After acquiring (§4.1) and qualitatively analyzing (§4.2) our sample, we conducted a computationally-assisted qualitative analysis of data visualization lineages (§4.3), in which we augmented existing data using digital forensics to trace the history and evolution of a subset of visualizations.

4.1 Data Collection and Sampling

We structured a computationally-assisted data collection protocol (Fig 2) to attain 100 data visualizations from X about immigration and/or the border crisis for qualitative analysis. A subset of these visualizations became our data visualization lineages, described in §4.3.

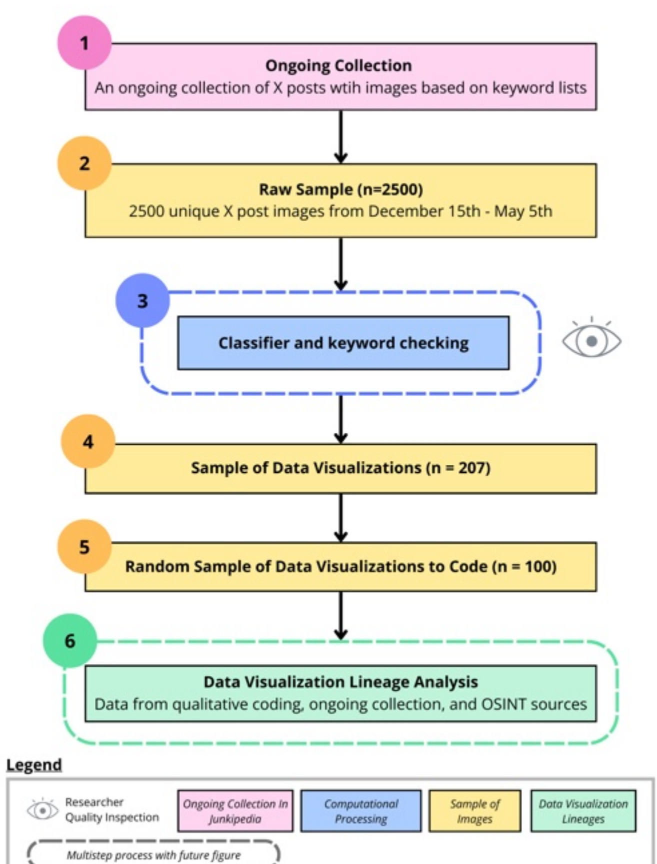

Fig. 2. Our process of gathering our 100 image random sample from a larger dataset, borrowed from an ongoing collection and computationally filtered by machine classification to support qualitative analysis of coding and data visualization lineages.

4.1.1 Social Media Data Sourcing. We collected X image posts from Junkipedia[1] using a set of keywords related to the border crisis and immigration. We used a grounded approach [53, 54, 159] to generate the initial list, first searching "border crisis" on X and then expanding our sample to include keywords and themes in the top results until 30 key terms, such as "open borders" and "sanctuary cities" were identified (listed in Appendix A).

Collection of this data was conducted from December 15, 2023 through May 15, 2024, and analysis occurred concurrently as part of a larger project. Borrowing from methodologies of theoretical and purposive sampling [53, 54, 135], the keyword tracking list was iteratively updated to include emergent relevant terms. We selected X due to its persistent salience in the broader political discourse [72, 159] and noteworthy cases of negative and misleading claims targeting asylum seekers and those biased against them [17, 18, 79].

We limited our collection to publicly available content with at least 1,000 views. This threshold helped protect contextual integrity by excluding low-view content that may not have been intended for public exposure and avoided amplifying fringe content by focusing on information already receiving attention [139, 156].

4.1.2 Sampling and Classification. To make it tractable for intensive human coding, where each piece of content could take up to ten minutes to qualitatively analyze, and to limit researchers' exposure to potentially harmful content, we selected a sample of 100 posts for coding. We summarize the steps taken to reach this sample in Fig 2.

First, we took a random sample of 2500 unique images from our broader collection. We computationally checked each image against one another for uniqueness using a structural similarity measurement index (SSIM) [23, 173] with a threshold of 0.95. We then ran this 2500-image sample through a classification and keyword-checking process (Box 3 in Fig 2), described in Fig 3. This yielded 207 computationally identified data visualizations (Box 4 in Fig 2).

We developed a classifier that assigned a probability (0-1) to images, indicating the likelihood of being a data visualization, with higher values reflecting greater confidence. Images with a probability of 0.85 or higher were inserted into the data visualization sample (Box 4 in Fig 2). Images within a threshold of 0.6 to 0.85 were checked for keywords to determine whether or not they were relevant data visualizations. Optical character recognition (OCR) was used to surface the text from the images and test these words against a list of keywords (such as "Percent", "Number of Crossings", "Number of Apprehensions"). These keywords were gathered inductively by manual examination of our larger dataset.

We trained our classifier with author-generated and free-use image training data from three online collections: Kaggle, Adobe, and Roboflow. In Table 1, we show a breakdown of our defined categories and training set coverage of different categories of data visualizations. For map-based visualizations, we used Python to generate a 5000-image dataset for the classifier, creating maps based on an even distribution of permutations of geographic characteristics ("Map Geography" in Table 2), randomly assigning image sizes in 16:9 and 9:16 ratios at 1080p resolution. Categories of data visualizations and map characteristics were guided by our research questions and established definitions in literature [92, 122].

After empirical testing, our classifier had a successful rate of ~88% accuracy through 10 runs and manual checking of our data. From the 207 images determined to be data visualizations from our 2500 random sample, we randomly selected 100 for qualitative analysis, resulting in our final "Random Sample of Data Visualizations" (Box 5 in Fig 2).

[1]Junkipedia (www.junkipedia.org) is an open source tool for collecting and monitoring social media and is a project of the National Conference on Citizenship (NCOC).

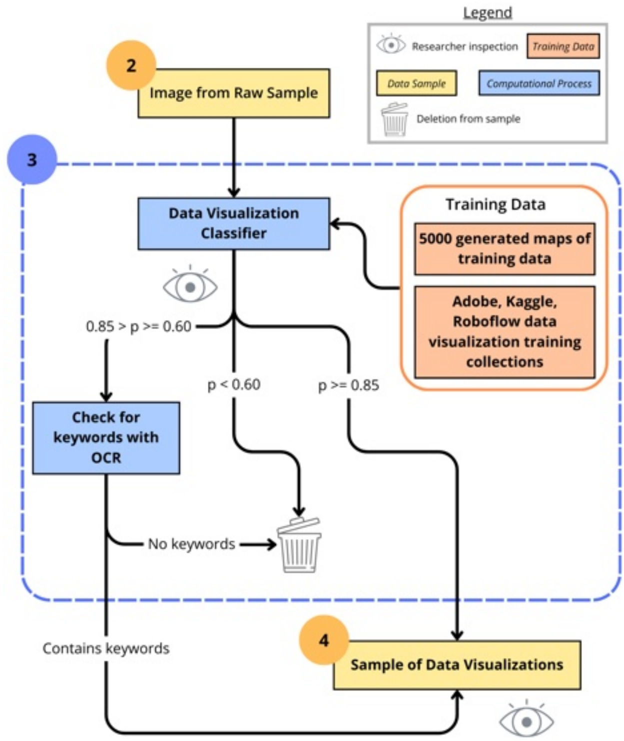

Fig. 3. Process by which images were determined to be data visualizations from our image classifier and keyword heuristics.

Table 1. Data visualization category coverage in image cluster training dataset.

Data visualization category	Training Data Sources
Bar chart	Kaggle, Adobe, Roboflow
Line chart	Kaggle, Adobe, Roboflow
Pie chart	Kaggle, Adobe, Roboflow
Sankey chart	Kaggle, Roboflow
Table	Kaggle
Scatter plot	Adobe, Roboflow
Donut	Adobe, Roboflow, Kaggle
Area charts	Roboflow
Maps	Self-generated in Python and GoogleMaps

Table 2. A breakdown of defined parameters for our maps.

Map Geography	Map Style
United States	Country outline
United States and Mexico	State outline
Mexico	Choropleth
United States and Canada	Grid cartogram (square)
Americas	Grid cartogram (hexagon)

4.2 Qualitative coding and thematic analysis

We applied thematic analysis, as specified by Braun and Clarke [41, 42], to analyze our 100-picture sample. First, the two first co-authors qualitatively coded all data. The codes in the codebook were derived deductively from literature (data visualization types [116, 122] and anti-immigrant rhetoric, [22, 38, 195]) and inductively [41, 42] from examining a larger collection of images (Box 4 in Fig 2) for key themes and trends.

Table 3. Code categories and definitions.

Frame Alignment	Data Units	Visual Characteristics	Visual Characteristics Definitions and Options
Presence of explicit anti-immigrant frame	X-axis units	Visualization type	Chart, Map, Infographic, Composite, Other
Counter to anti-immigrant frame	Y-axis units	Map type	International, Country, State, City, Other, Not a map
	Other units (if not on axes)	Chart type	Bar, Line, Area, Sankey, Pie, Other, Not a chart
		Visualization source	Cited creator or source of the visualization
		Data source	Cited source(s) of the data visualized
		Annotation	If a post-production annotation was added
		Watermark/Logo	If a watermark or logo is visible

We coded visual characteristics of the data visualizations, such as the type of visualization, named sources, and logos or watermarks. We also coded data units of the visualizations and for frame alignment if a visualization aligned with or countered an anti-immigrant rhetorical frame. We outline our codes in Table 3. The co-authors met to perform IRR and come to consensus. We calculated an average Cohen's Kappa of 0.746 [56] with a minimum of 0.681 and maximum of 0.791 across our codes. Individual code breakdowns are in Appendix C.

After coding we conducted thematic analysis [41, 42]. The first round was to split the sample by chart type and then by groups of codes, such as key data units, particular anti-immigrant rhetorical frames, and visual annotation choices. Rhetorical frames were identified through themes in the data visualizations and using Entman's model of framing [81], focusing on how the visualizations emphasized certain issues, such as moral judgments or diagnosing societal problems towards immigrants. In our axial coding round, we identified themes from these thematic and code groups, described in §5. During this analysis, we discovered related data visualizations, which we define as "Data Visualization Lineages". These lineages necessitated a computationally-assisted qualitative analysis to assemble and analyze them in depth.

4.3 Computationally-Assisted Qualitative Lineage Analysis

We performed a detailed analysis of three data visualization lineages. For each lineage, we utilized SSIM to find related images in our larger database (Process B in Fig 4) and leveraged Open Source Intelligence (OSINT) tools to fill in gaps from our data collection (Process C in Fig 4). We captured this investigation in timelines in a digital whiteboard software, which we qualitatively analyzed for design choices made between different iterations and online dynamics of spread, described in §5. We summarize this process in Fig 4.

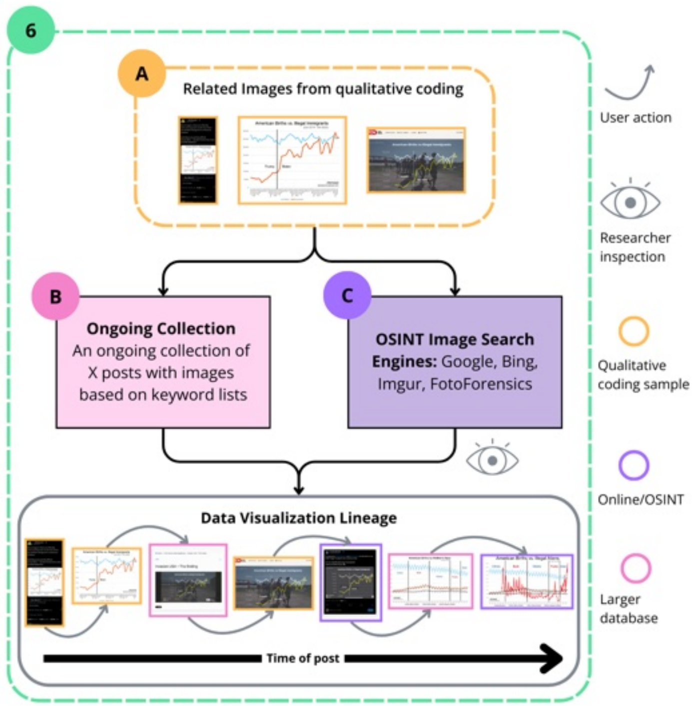

Fig. 4. Related images came from qualitative coding, the database with an ongoing collection, and OSINT technologies to form data visualization lineages.

In Process B (Fig 4), we compared these related images to images in our larger, ongoing collection. We did this comparison with strategic filtering based on keywords in the visualizations via OCR (as done in §4.1.2), matching colors and font faces, and SSIM (also as in §4.1.2) [23, 126, 173]. From this, one of the co-authors reviewed the resulting images and added these to the lineage (Fig 4). In Process C (Fig 4), we leveraged OSINT tools and methods to understand more images provenances and occurrences in the wider internet [9, 98, 117]. We specifically leveraged reverse image search engines (Google, Bing, Imgur) to inform us about spread and OSINT tool FotoForensics aided in detecting image manipulation to add additional context in user remixing behavior in our lineages.

With results from both processes, we put together our lineages as not just a series of chronological posts but with assigned user actions taken, represented by arrows in Fig 4. User actions were defined through qualitative analysis by two authors and literature [122, 179]. Chronological order of posting was determined by metadata from Junkipedia or Process C investigation. We describe these lineages, tactics, and online dynamics in which they emerge in §5.4.

4.4 Reflexivity and Positionality

As with all qualitative work, we acknowledge that it is impossible to remain completely neutral, agnostic, and detached from our study topic. Our team shares diverse experiences shaped by their various identities, values, and political leanings that influence our approach to this topic. We draw on qualitative traditions for fostering reflexivity in this research [25, 111, 120] as we are “human research instruments” [87] studying political discourse and propaganda about immigration in the US—a real time, contentious issue that is ever evolving. Although our team members have different perspectives and political alignments, none of our political beliefs share the anti-immigration framings that were prominent in much of our data. As such, the analyses we develop in this paper have been shaped not only by our research lenses, but also by our personal and political ones.

5 Findings

Below we describe the overall composition (§5.1, §5.2) and key themes of our sample (§5.3). Alongside this thematic analysis, we introduce and illustrate our notion of Data Visualization Lineages through three case studies (§5.4).

5.1 Types of Data Visualizations

Informed by literature [49] and our image classification data (§4.1.2), we observed five primary groups of data visualizations in our sample: Charts, Maps, Composite charts, Infographics, and Other. We define Charts as any type of information shown in a graph and Maps as any geographic representation of information. Meanwhile, Infographics are images showing statistics, numerical claims or immigration-related processes. We define Composite charts as images with multiple charts or as charts superimposed onto one another or other infographics (see Fig 5). This composite chart definition is similar to previous work considering images and infographics with multiple charts [74, 122]. We used this distinction to count images that may have been styled as multiple charts but were still units consistent as one chart for ease of analysis at scale. We use Other as a catch all code for visualizations that did not fit the other categories.

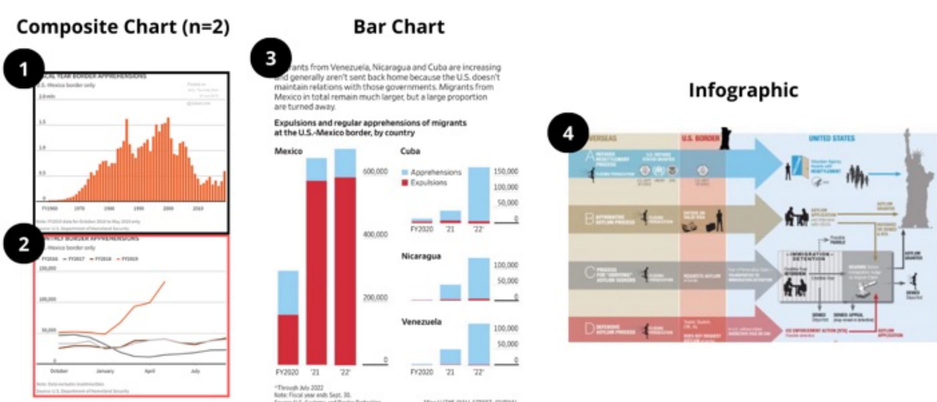

Fig. 5. An example of an image counted as a composite chart and being broken into 2 images (Image 1 and Image 2) – one an image of a bar chart, another of a line chart. Meanwhile, Image 3 is counted as one bar chart given this is a stylistic breakdown of comparable units that could have been on the same chart. Image 4 is an infographic – presenting an immigration process without traditional chart axes or data units.

Our 100 item sample included 88 visualizations that were relevant to the immigration crisis, including five distinct visualizations spread across two composite charts. Table 4 shows that charts dominate this sample at 76.5%, with bar and line charts as the most prominent (Table 5). Due to this overwhelming prevalence, we focus on charts in our analysis.

Table 4. Distribution of types of visualizations in our coded visualizations.

Type of visualization	Chart	Map	Composite Charts	Infographic	Other
Frequency in Sample	76.5%	15.2%	3.5%	2.4%	2.4%
Count	66	13	2	2	2

Table 5. Distribution of chart types in the charts subcollection.

Type of chart	Bar	Line	Area	Sankey	Pie	Other
Frequency	56.3%	28.2%	5.6%	1.4%	1.4%	7%
Count (of 71)	40	20	4	1	1	5

5.2 Summary of Data Unit and Source Trends

We present our data sample's composition, focusing on data visualization units, source attribution trends, and how these intersect with creators' choices to support specific rhetorical frames.

5.2.1 Data Visualization Units. Many of the sampled visualizations (~46%) use time as a unit of measurement. Of these, 54.3% measured "Apprehensions", "Encounters", and "Crossings" — units adopted from US Customs and Border Protection (CBP) open online data [4]. In total, ~28% of our entire sample use units of "Expulsions", "Encounters", "Apprehensions", and "Crossings" and in the visualization cite CBP as the source of these units. Another CBP-cited unit, "Inadmissibles", is far less common (n = 1). We display CBP official data units from their data glossary [6] in Table 6.

Table 6. Governmental Immigration Units from the US Customs and Border Protection (CBP) Datasets [6].

Term	**Definition from Homeland Security**
Expulsion	Noncitizens expelled under Title 42 authority to their country of last transit or, if a person cannot be returned to the country of last transit, to their country of origin.
Encounter	The sum of U.S. Border Patrol (USBP) Title 8 apprehensions, Office of Field Operations (OFO) Title 8 inadmissibles, and noncitizens processed for expulsions under Title 42 authority by USBP or OFO.
Apprehension	The arrest of a potentially removable noncitizen by the Department of Homeland Security (DHS).
Inadmissibles	Persons not allowed to enter the United States due to a variety of reasons such as health, crime, national security, or fraudulent documentation.

These CBP units have distinct governmental definitions and sources. In our sample, they are frequently renamed into a non-CBP label entitled "Crossings" [6] (Image 1 in Fig 6), not being attributed to CBP (Image 2 in Fig 6), or used interchangeably in charts (Image 3 in Fig 6). We often find this to be misleading, as many of these units do not represent migration but rather individuals that are stopped at the border and may never enter the US.

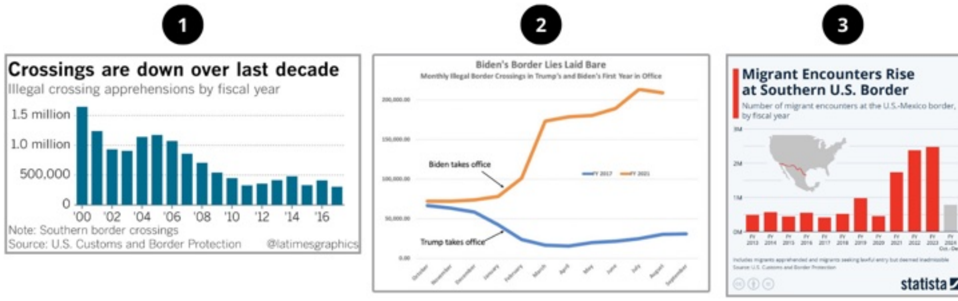

Fig. 6. Illustrating 3 key tactics and examples of CBP data units in our sample. Image 1 shows an example of "Crossings" as a category standing in for "Apprehensions". Image 2 shows no citation to CBP data. Image 3 shows "Encounters" being combined with "Apprehensions".

Some visualizations (Image 2, Fig 6) do not cite CBP data which makes it difficult to understand the origin and accuracy of these trends. Other images misleadingly combine data series or use these data units interchangeably, (Image 3, Fig 6) where we see a chart representing "Encounters". However, text below this chart says it is actually combining units of "Apprehensions" and "Inadmissibles", which is misleading in that these are different data units and it is unclear how "Encounters", "Apprehensions", and "Inadmissibles" are being combined in this chart.

Although CBP is the most frequently cited agency, other government agencies also provide public data used in migration discussions. A key rhetorical frame in our sample is demographic shifts (§5.3.1), where US Census and CDC natality data (tracking population growth via birth rates) are cited in ~8% of charts in our sample.

5.2.2 Sources of Data and Data Visualizations Are Not Often Attributed. For each visualization in our dataset, we qualitatively coded both the cited source(s) of the visualization and the source(s) of its underlying data (for example, the cited source of Image 3 in Fig 6 is Statista, while the data is cited as CBP). In both cases, source citations were often lacking—with 39.2% of our dataset missing citations for the visualization and/or data sources. The most cited data source, 22.8% of our sample, was the CBP's public online database [4]. The second most cited data source was Pew Research Center, a nonpartisan think tank [154], at 3.9%. This illustrates just how large the spread of this sample was across different data sources. Of the visualizations with cited individual creators, the top three, all at 5.1%, were: Pew Research, Axios Visuals, and our first lineage's content creator. We provide a full breakdown of the data visualization and data sources in the Appendix B.

The ability of creators and later remixers to remove identifiers of origins of the data and visualizations (which occurred in 39.2% of our sample) highlights the challenge of tracking the provenance of visualizations. We explore this concept further in our case studies on data visualization lineages introduced in §5.4.

5.3 Anti-immigrant Frames Present in Data Visualizations

We find four salient rhetorical frames from our qualitative analysis: Demographic shifts (§5.3.1), Biden-Harris administration culpability (§5.3.2), Criminality in immigrant populations (§5.3.3), and Immigrants taking welfare from taxpayers (§5.3.4). To describe these frames, we use Entman's framing theory [81] and key functions, due to their focus on media objects and political communication while providing an operationalized framework to describe and evaluate specific choices made by data visualization creators that support or counter different frames. However, not all frames or visualizations include every element of Entman's schema.

5.3.1 Demographic Shifts Due to Immigration. The data visualizations support a frame of demographic shifts by leveraging population data in two key ways: 1) comparing demographics over time (Image 1, Fig 7) or physical space (Fig 9), and 2) juxtaposing immigrant population data with domestic population data (Image 2, Fig 7, and Fig 8).

These visualizations commonly cite census or natality data from the CDC. Small changes in data visualizations can invoke specific subframes within demographic changes. Two subframes of note are the Great Replacement Theory and an adversarial view of demographic shifts, with the first framing immigrants of color as a threat to white populations and the second framing opposition between domestic and immigrant non-white populations.

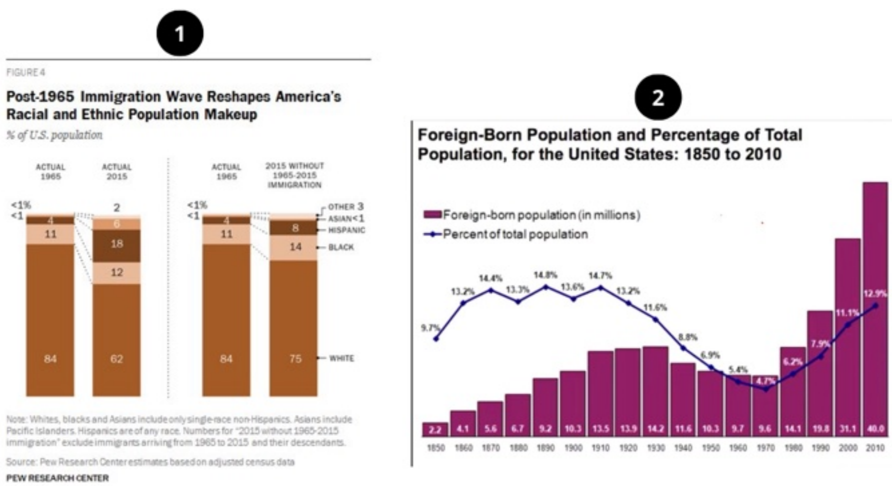

Fig. 7. Examples of visualizations in the broader subframe of demographic shift. Image 1 shows racial demographics over time in the US while Image 2 pairs natality population data with immigration data. These represent two overarching tactics used in this frame.

The Great Replacement Theory (GRT) The Great Replacement Theory (GRT) is a white supremacist conspiracy theory that is a specific instantiation of other replacement theories, as analyzed by Bracke and Aguilar [40] and others [90, 141]. This theory, which has gained popularity in mainstream social media conversations [71, 109], claims white, Western populations are being systematically replaced by migration from majority non-WEIRD countries [2, 40, 90]. We see this theory emerge as a subframe in our dataset.

Visualizations in our dataset reinforce this subframe by dividing racialized and immigrant natality data (Fig 8, §5.4.1), producing evidence that white populations are being outnumbered, or replaced, by non-white immigrants and their descendants. Fig 8 illustrates the division between non-Western immigrant natality data and their descendants compared to white Danish natality data. This generational component of natality data could be nefariously interpreted as a systematic population replacement effort, placing moral judgment on non-white pregnant immigrants in Denmark. Similar themes appear in the US, explored in our first data visualization lineage (§5.4.1).

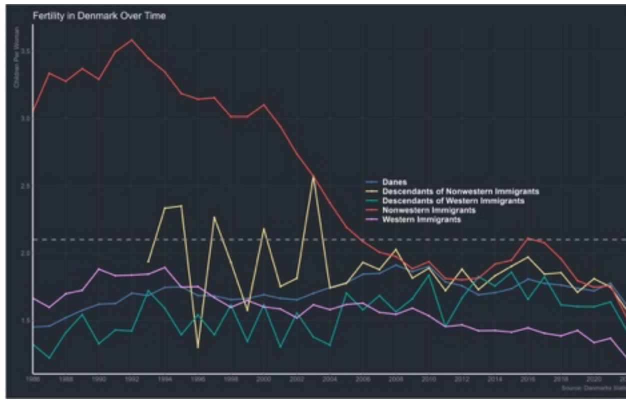

Fig. 8. Denmark natality data, separated out by immigrant and non immigrant populations, divided out by groups of "Western" and "Nonwestern" immigrants and their descendants. This combination of racialized immigration and natality data with a "Western" and "Nonwestern" divide provides evidence for GRT.

Adversarial Demographic Shifts

Our dataset also reveals rhetoric that pits non-white populations against each other in the context of recent mass migration. This discourse is often economic, focusing on jobs[2] and resource access. Fig 9 shows how a map of high immigrant populations is framed as evidence of a government effort to lower wages in Black communities by employing immigrants.

Fig. 9. A map depicting a subframe for demographic shifts pitting marginalized American groups (in this case, Black Americans) against immigrant groups.

5.3.2 Biden-Harris Administration's Culpability in Uncontrollable Illegal Immigration. This frame focuses primarily on diagnostic causes, ascribing blame to the Biden-Harris administration's handling of the US-Mexico border as the cause for unprecedented migration. In some visualizations, the Biden administration immigration numbers are juxtaposed against the first Trump administration's; suggesting a second Trump term could remedy uncontrollable illegal immigration.

[2]Our data collection ended before the "Black Jobs" discourse after the Presidential debate [108], but we find the existence of this rhetoric before this popular discourse about "Black Jobs" as indicative of the pervasiveness of it.

We find this framing primarily enacted through 1) supportive annotations adding "Biden" and/or "Trump" labels to visualizations [121, 122], 2) timeline extensions in visualizations that accentuate the increase in land border encounters under Biden compared to past administrations (namely Trump), and 3) explicit use of chart titles that blame the Biden-Harris administration. These tactics are often used in tandem with one another (examples in Fig 10).

Some of these tactics selectively add historical context via annotations of presidential terms and policies and also leave out sociopolitical context (immigration during COVID-19's peak and Trump's term in 2020 was far lower across the world [88]). We also observe in §5.2.1 that this frame is often supported by misleading usage of US government migration data.

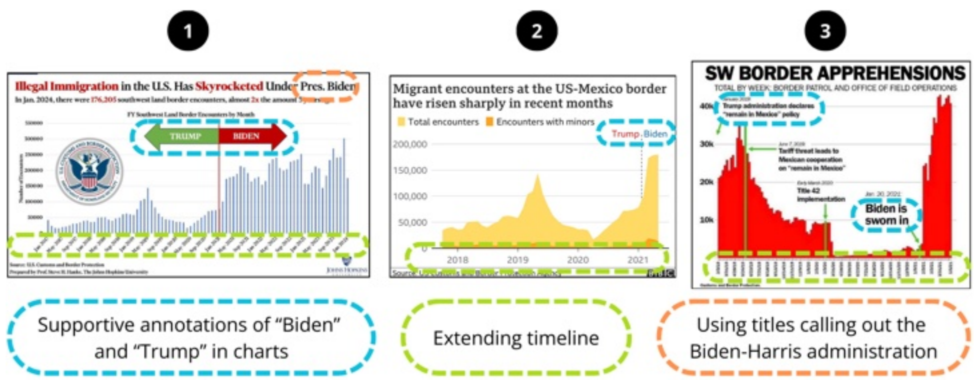

Fig. 10. Charts showing immigration numbers under different presidential administrations, often with supportive annotations, extending the timeline to include more administrations without external context, and titles explicitly calling out the Biden-Harris administration.

Visualizations countering this frame use "Apprehension" data (Table 6) to challenge the "open border" narrative (see Fig 11), trying to illustrate enforcement against illegal immigration. However, since apprehensions do not guarantee removals and rising immigration may naturally lead to more apprehensions, such visualizations can be misleading.

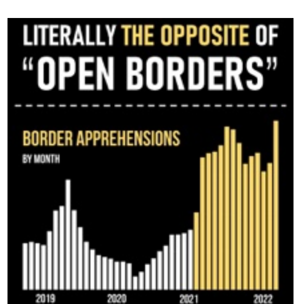

Fig. 11. Countering the "open borders" narrative by depicting increases in apprehensions. This chart is misleading due to apprehensions guaranteeing removals and not acknowledging that increased immigration may increase apprehensions.

5.3.3 Misleading Claims of Criminality in Immigrant Populations. A key frame in our data visualizations claims that immigrants entering the US through the Southern border are predominantly criminals ("murderers" and "rapists") who would introduce violent crime into majority-white

American communities. This frame focuses on how "these countries" (majority non-white nations) have high rates of violent crime and are "sending their worst", specifically criminal non-white men, creating a strong moral judgment against these populations. We see this in action in Fig 12, which utilizes CBP "Encounters" units juxtaposed with reported rates of murderers and rapists per 100k capita in different countries to extrapolate numbers of violent criminals coming into the United States. This association is misleading, as it is not statistically accurate to apply a national-level population crime statistic to a subpopulation of immigrants. It is also unclear how this statistic was calculated. Other tactics include supportive annotations and titling of charts to invoke a frame of increased criminality in immigrant populations (§5.4.2). Furthermore, this criminality association creates a problematic view of immigrant groups supported by seemingly credible evidence through this combination of different data.

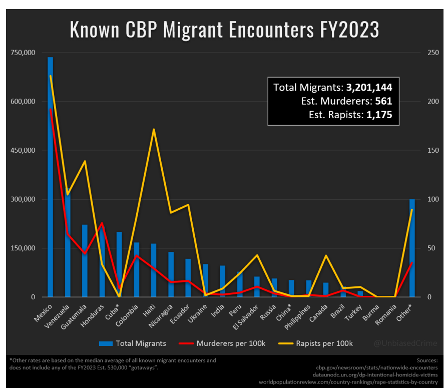

Fig. 12. A graph showing migration volume by country with reported criminality by country, aligning with a criminality framing with supporting annotations of "estimated rapists" and "estimated murderers".

This type of criminality framing has led to dangerous offline violence in European nations [99] and anti-immigrant policies [16], and also counters evidence that immigrant populations often commit less violent crime in the United States than US-born individuals [28, 125, 147]. In our data, this frame is countered by such statistics, such as Fig 13. This composite chart also compares similar units as Fig 12, using 100k per capita average for all comparisons. However, it is worth noting this chart does not have clear data citations.

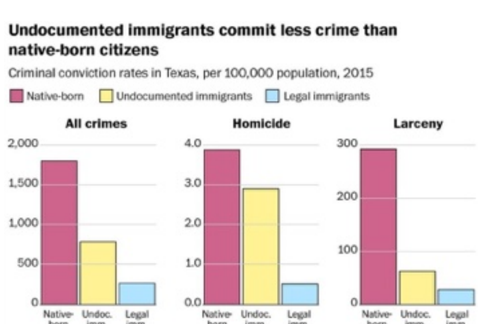

Fig. 13. Countering a criminality narrative showing immigrants commit less crime than citizens.

5.3.4 Immigrants Taking Welfare From Taxpayers. This frame highlights social services (for example shelter, food stamps) received by immigrants, often labeled as "welfare" [13] and portrays them as resources taken from taxpayers and citizen populations. This frame builds on discourse of non-citizens receiving welfare [12], diagnosing the cause of high taxes and strains on overburdened welfare programs as immigrants coming for "free handouts" at the expense of US taxpayers. We often see this frame supported through maps which show costs per state of immigration or in infographics that cite budget cuts to services such as education and hospitals to fund migration services (Fig 14). Importantly, these strategies are misleading because these budgets vary across states and federal line items [5, 10], and verifying these claims is difficult.

Fig. 14. A map and infographic stressing a welfare narrative by showing tax estimates by state and claiming the amount of money taken from other budgets due to immigration.

5.4 Data Visualization Lineages

Here, we highlight and unravel Data Visualization Lineages (DVLs): families of data visualizations that share the same origin but diverged through distinct visual alterations enacted with technical affordances (screenshots, cropping, photo-editing, annotations) [122]. DVLs connect the notion of image provenance and participatory visual culture to data visualizations that circulate and evolve online. We show how visual changes are enacted to shift meanings (to resonate with different frames) and to obscure the provenance of the original visualization and its underlying data.

We present three case studies of data visualization lineages which support different frames and have distinct styles and sources, but all cite publicly available government data about immigration and natality. For each lineage, we describe the results of our qualitative analysis of the sample from the lineage acquired from the methods in §4.3, acknowledging that our sample does not encompass the entire lineage online and that these lineages are often branching and growing in real time. The first lineage (§5.4.1) features a misleading chart about natality and demographic shifts, spanning 15 analyzed iterations. The second features a longtime series of data visualizations initially used by a US Senator, then later turned into the Trump campaign media piece highlighted in this paper's introduction (§5.4.2), encompassing 28 unique analyzed visualizations and images of them. The third focuses on visualizations generated via screenshots of the CBP government dashboard, which we estimate to be the largest branching lineage of all of our data. We analyzed 21 images of a particular event—examining how users create, remix, and misappropriate evidence from these dashboards to fit their frames (§5.4.3).

5.4.1 Viral, Misleading Natality and Demographic Shift Claims. Our first DVL is a four-month lineage of a correlational line chart linking domestic natality rates to immigration, employing Great Replacement framing (§5.3.1). Promoted by Elon Musk on X, this chart went viral, sparking iterations (Event 1). Media personality Tucker Carlson later repackaged and amplified this chart across platforms and in a documentary (Event 2), while the original creator produced content around this frame for months afterwards (Event 3). Musk's promotion and the chart's cross-platform spread highlight how visualizations can shape, and be shaped by, political conversations on online platforms. Fig 15 summarizes this lineage, showing key events, user iterations, and origins of our lineage images.

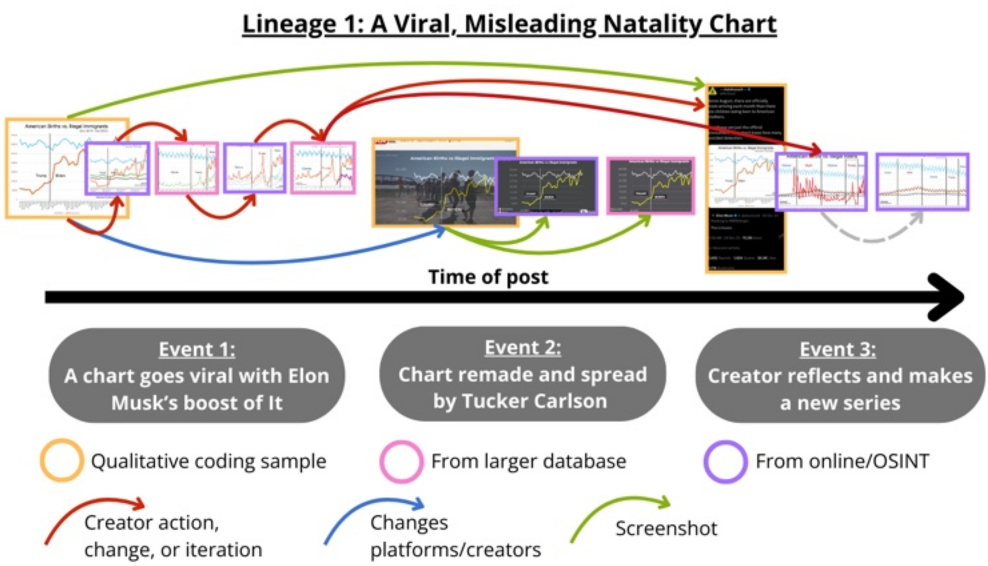

Fig. 15. The data visualization lineage for this viral natality graph, divided into three distinct events and showing user actions at each stage.

Event 1: A chart goes viral

The lineage begins in December 2023 from a creator who regularly posts charts about crime, immigration, welfare, and other US sociopolitical topics. At the time of writing, this verified account has over 136k followers on X. This creator posted a chart as a reply to Elon Musk, owner of X, who had commented on a thread about the US-Mexico border. The line chart, titled "American Births vs Illegal Immigrants", shows two plotted series: 1) number of births in the United States, and 2) number of Encounters at the US-Mexico border (titled "Illegal Immigrants"). The US Center for Disease Control (CDC) for natality and US Customers and Border Patrol (CBP) for Encounters were respectively cited as the data sources in the thread (though not on the chart).

The equivalence of Encounters (defined in Table 6) to a total number of "illegal immigrants" is false. The spurious and misleading juxtaposition of this metric with natality data underscores the Great Replacement frame (§5.3.1). To support the frame of the Biden administration's culpability, the creator placed an annotation line at January 2021 to separate Trump and Biden's terms, a common annotation seen in §5.3.2.

Within an hour, Musk shared and pinned the post (Fig 16). In doing so, Musk boosted the chart to a massive audience where it received millions of views (76.5M at time of writing). Subsequently it garnered mass media coverage [43, 109] and in the following weeks the creator posted several iterations of the chart growing the lineage. This demonstrates a "post-spotlight" effect observed in prior studies where creators receiving significant new attention [165, 199] adapt their content to the perceived appetites of their audiences.

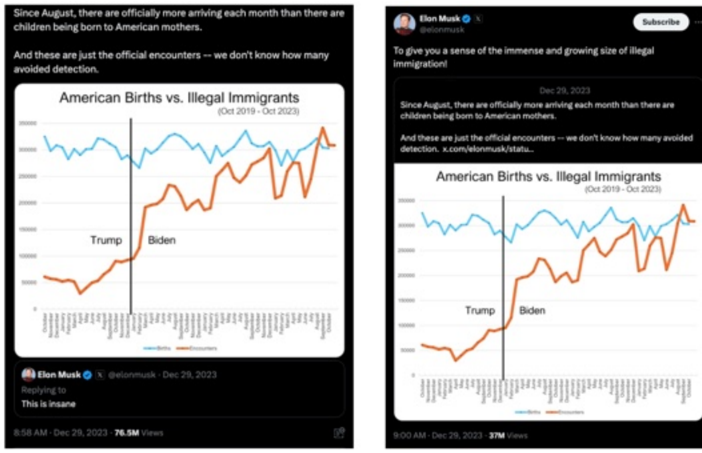

Fig. 16. Showing the creator's original post and Musk's reposting and pinning of the chart. These images have been cropped or edited for anonymization.

Post-spotlight iterations of the chart (summarized in Fig 17) focused on adding new data series, such as adding more publicly available CBP data units in one chart (Image 3, 4, 5) or comparisons between different racialized natality data series (Image 2), a tactic seen in other Great Replacement charts in our dataset. Another tactic was extending the x-axis (Images 3, 4, and 5) to elongate the historical lookback period, as seen in §5.3.2 to stress the Biden administration's culpability.

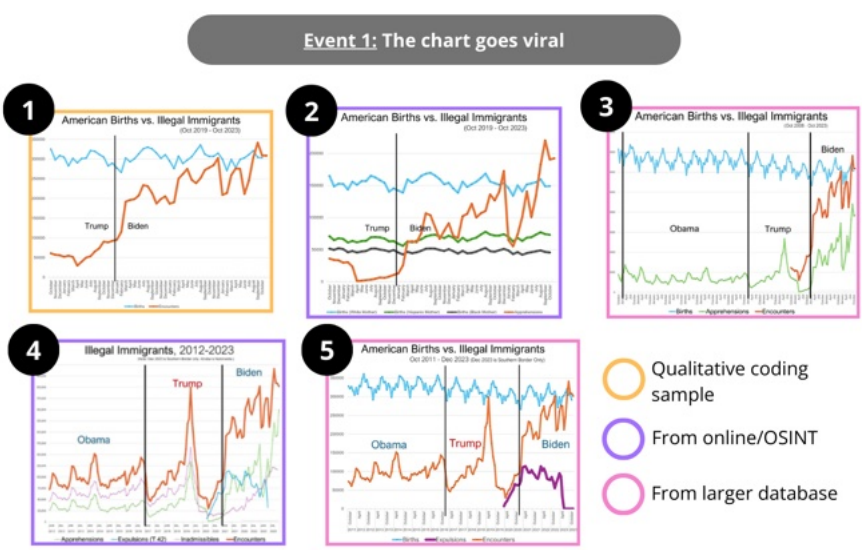

Fig. 17. Descendant charts of this lineage, showing the addition of President Obama's term years and new data series. Images have been edited for anonymization purposes.

Event 2: The chart crosses platforms and creators

On January 17th, only 2 weeks after the lineage began, Tucker Carlson's media company, Tucker Carlson Network (TCN), used an animated version of the initial chart in a documentary published on YouTube and their website. Media outlets reported on this chart and Carlson's use of it, and the TCN video was flagged on Facebook for being misleading [21].

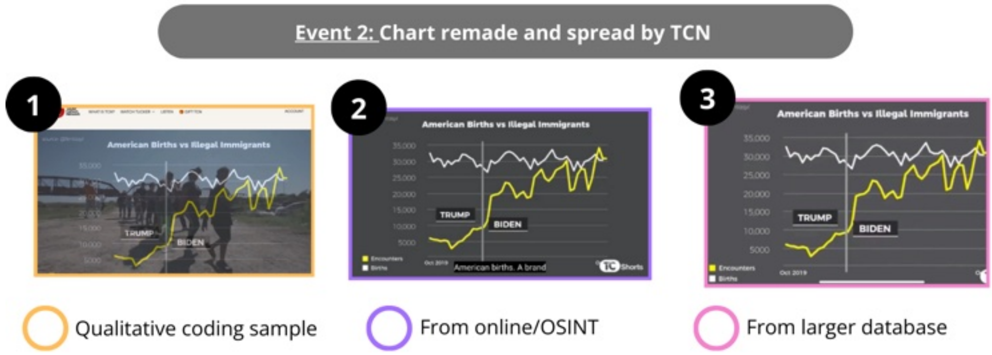

Fig. 18. A large amount of images of the chart are created due to the animated background in the TCN film, causing significant spread of this chart and slight variations in this point of the lineage. Here we see 3 examples of different aspect ratios and video affordances of the background of the chart, the TCN website's portal (Image 1) and even closed captions on Image 2.

The re-creation of the chart into an animated sequence produced numerous similar but slightly different images, as the animated background was captured at different moments and screen ratios (Fig 18). This marked a significant shift from the earlier spread of static image uploads from the creator and amplified the content beyond X and through the TCN media ecosystem.

Event 3: The chart gets a new series

In March, the creator acknowledged how the boost of the first chart impacted their following and internet presence. In April, they continued to create more charts expanding this lineage. These charts use their initial tactics of iteration in adding new data series and extending the x-axis into previous presidential administrations, illustrated in Fig 19.

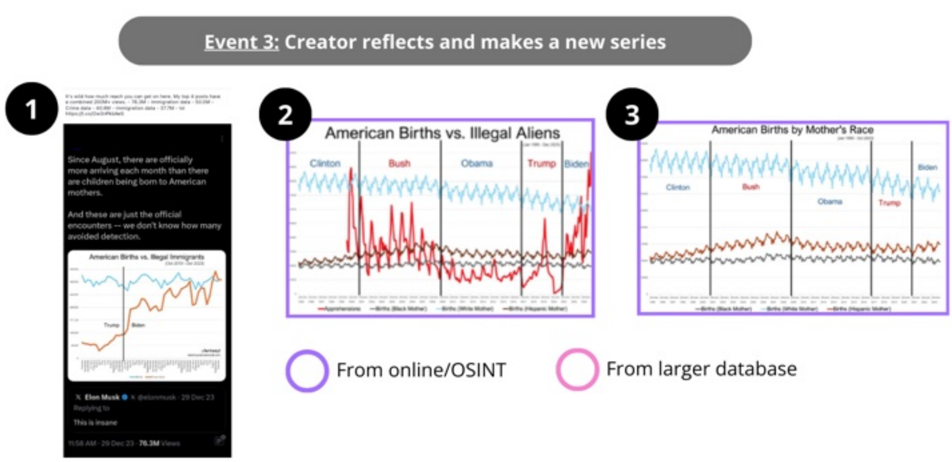

Fig. 19. The creator reflects on the spread of the initial post and continues this lineage in April by adding more Presidential terms and racial natality data. Image 1 from Junkipedia.

This lineage demonstrates how spotlighting a chart can drive iterative creations of visualizations that support prevailing political frames (in this case, the Great Replacement Theory) and spread across networks, mediums, and audiences. In this lineage and others, we also see how easily visualizations of public government data can be created in response to current events and support new anti-immigrant rumors.

5.4.2 Longtime Senate Visual Aid Turned Trump Rally Centerpiece. The second lineage comes from the office of Wisconsin Senator Ron Johnson, who used CBP data to create visualizations as illustrative aids to fit his political framing that the Biden administration had not properly governed the US Mexico border, leading to record-breaking immigration and safety issues, a frame explored in §5.3.2. This lineage is co-opted by the Trump Presidential campaign, who created a massively popular chart from a photo-edited version of Johnson's. This lineage, summarized in Fig 20, is unique in its presence in physical and offline spaces, design iterations by the Senator's office, and increased spread via Trump's repeated promotion of an image within this lineage as part of his Presidential campaign.

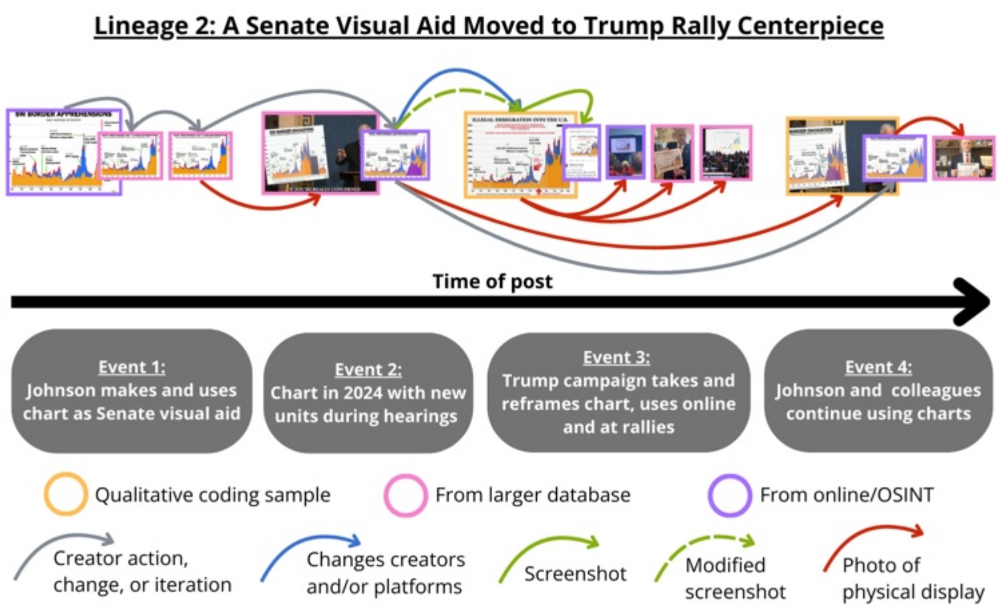

Fig. 20. Summary of the second data lineage showing a chart made by Senator Johnson evolving and then being co-opted by the Trump campaign for use in rallies.

Event 1: Johnson's office create and iterate on the original chart

In March 2021, Senator Ron Johnson's office released the first version of this chart (Fig 21). The Senator frequently used these charts in Senate meetings, press coverage, and on his official X account [44, 143]. These charts provided the Senator with a communication aid with his colleagues and audiences as well as a specific visual brand attached to his border security platform. Through 2021-2023, the chart was a stacked barchart of Southwest Border apprehensions separated into three demographics: "Unaccompanied Minors", "Families", and "Single Adults". Annotations of key policy decisions and public declarations were placed on the chart in their corresponding years.

Johnson utilized the official CBP data in his charts, but repeatedly iterated on the chart units, such as changing from daily to monthly totals for larger y-axis values (Fig 21). This change, a common strategy in manipulation of charts [122], amplified the framing of the border crisis by increasing the y-axis by an order of magnitude.

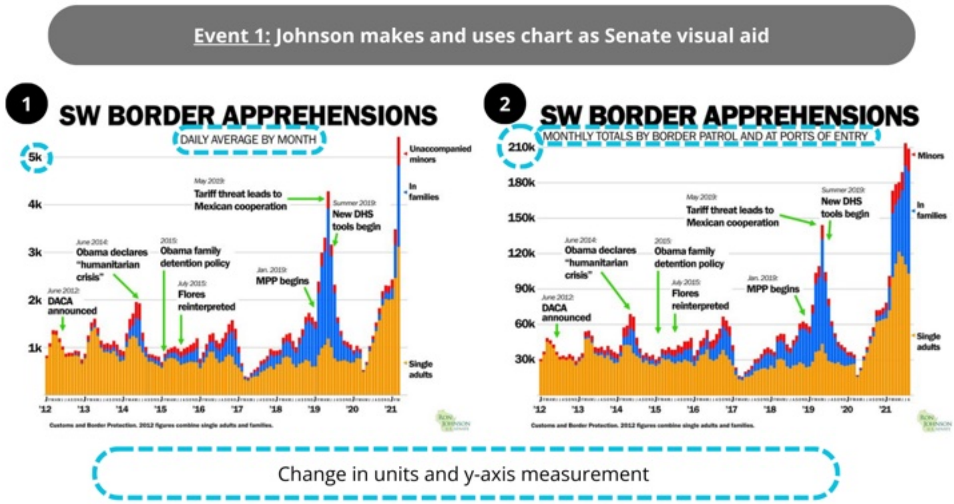

Fig. 21. Original March 2021 (Image 1) chart showing daily average by month encounters from January 2012 to January 2021 of three demographics, quickly changed to units of Image 2 for monthly encounters. This is the start of the lineage and the first iteration within it to produce visual evidence of a different scale.

Event 2: Newer versions of the chart appear in 2024 for key Senate hearings

Johnson, and other politicians, appeared with new versions of these charts again in January and February 2024 for a Senate hearing about a border bill and aid for Ukraine [7] (Fig 22). In this version, Johnson's office made a specific visual annotation (Fig 22) to compare Biden's migrant flow proposal to Obama's smaller 2013 numbers.

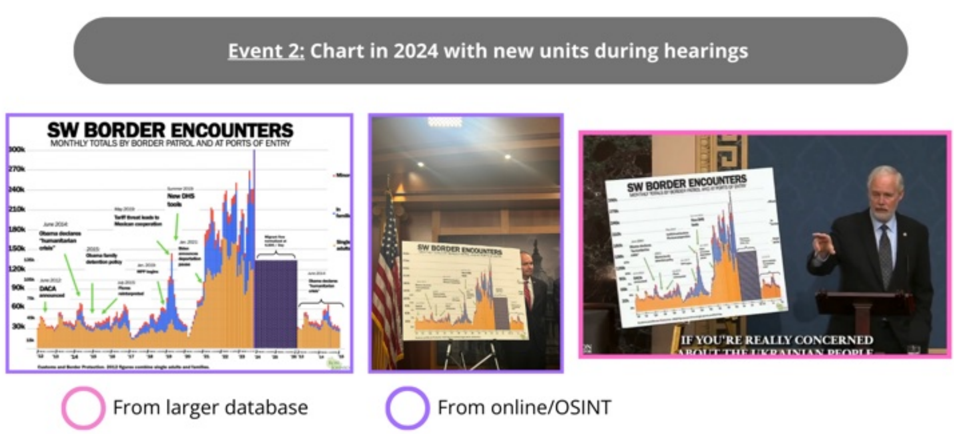

Fig. 22. Showing the January and February 2024 versions in the Ukraine hearings, as well as their origins in our data lineage collection.

By selectively annotating and cherrypicking data (2013 vs. 2024), the chart attributes border encounter growth to Biden's policies. Extending the x-axis further frames this increase as a policy-driven failure worse than other administrations. In March and April of 2024, Johnson released another chart (Fig 23) on X and in a press release during his and his colleagues' Senate hearing on impeaching Homeland Security Secretary Mayorkas [8]. This chart, styled like previous charts through 2021-2023, adds a new data series on Title 42 expulsions (in purple in Fig 23). This addition exposes a flaw in either this chart or its predecessor (Fig 23), as the totals remained unchanged, implying the earlier chart may have counted expulsions as "Single adults" rather than removals. This is misleading, implying high levels of entering immigrants, as Title 42 expulsions refer to removals during the COVID-19 public health emergency declaration in 2020 - 2023 [3].

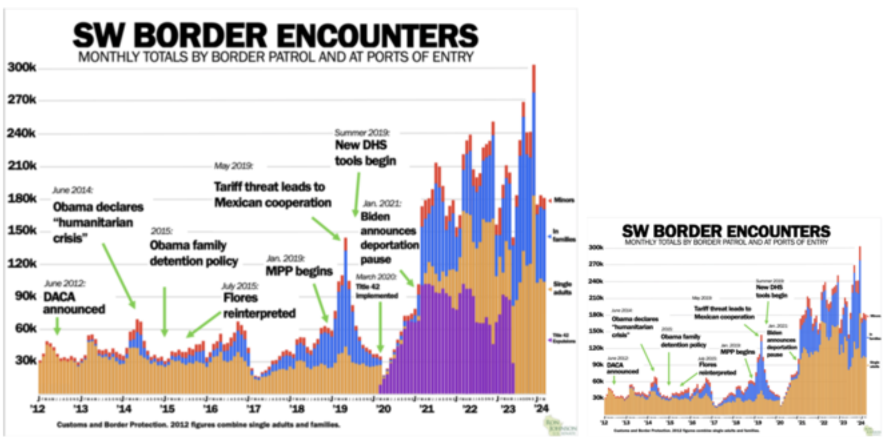

Fig. 23. A version of the chart with "Expulsions" added in purple (left). Smaller, original chart (right) does NOT include "Expulsions" (purple series) but maintains the same totals.

Johnson has been able to establish a brand of visual communication via data visualizations, where through annotations, changes of data series and time scales, he can produce evidence to support his political frames around immigration in his role as a Senator. This lineage explores the outcomes of his visual evidence and what occurs when these artifacts are co-opted by another politician for different framings.

Event 3: The Trump campaign edits and co-opts Johnson's chart

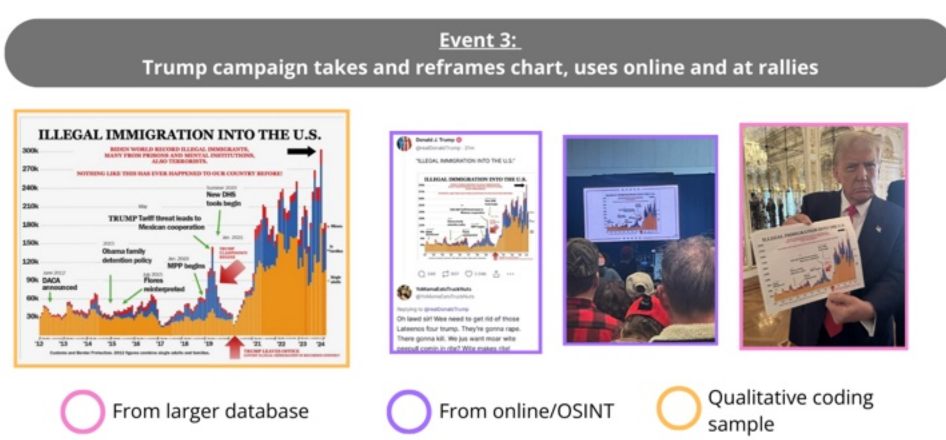

Fig. 24. Summarizing our image collection of this event in which a modified version of Johnson's chart appears from the Trump campaign and in photos with former President Trump.

This lineage shifted significantly on April 2, 2024, when Donald Trump posted an edited version of Johnson's March 2024 chart on Truth Social (Fig 24) [44]. This version used textual annotation edits (by removing, editing, and adding to Johnson's annotations), renaming of the chart, and photo editing of the data series to remove expulsion data, while also cropping away the data series legend. Johnson's logo was also removed and he was not credited in the chart (Fig 25).

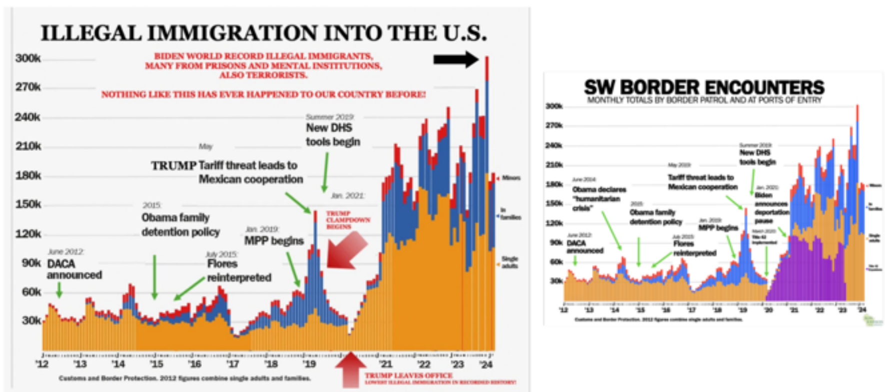

Fig. 25. Trump's chart version (left), with edits to Johnson's (right) by removing "Expulsions", adding new annotations, removing some key annotations, retitling the chart, and removing Johnson's logo.

These edits accentuated the frame of the Biden-Harris administration's culpability and Trump as a potential remedy to this immigration crisis with supportive annotations and the removal of Title 42 expulsions series. Furthermore, a frame of criminality (5.3.3) was added to this chart with titling and annotations that misleadingly implied that these spikes came from violent and dangerous immigrants. This chart rapidly grew popular across X and Truth Social. It was printed and displayed at a rally on April 2nd, the same day it was posted on Truth Social, and Trump was photographed with a print-out of it on April 6 (Fig 26). Later, in July, Trump used this chart at his rally when an assassination attempt occurred as he was pointing at this chart (Fig 26).

This lineage showcases not only how these charts circulate online, but also how they leave the digital space in both formal display materials for Senate Hearings and campaign events (Fig 22, 26). Politicians being photographed with these charts may be seen as an endorsement and emphasis of the frames within them, including Trump displaying himself as a remedy to the issue of "illegal immigration" framed in his version.

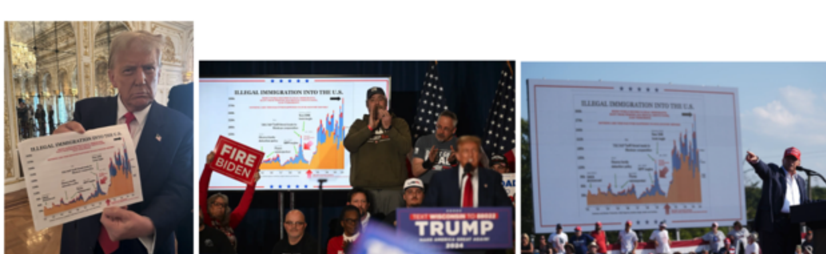

Fig. 26. Trump with a print out of a chart on April 6 (X user), with the chart on display at an April 2 rally (Joshua Lott/Washington Post), and at the rally of his assassination attempt on July 13 (AP Photo/Evan Vucci).

Despite some users on X pointing out the connection between the charts, no public comment has been released from Trump or Johnson regarding the chart. However, after Trump's co-option and photos with and of the chart, Johnson and other politicians continued sharing Johnson's versions in photographs and as visual aids.

This lineage illustrates how a Senator used CBP data to create visual evidence framing the Biden administration as responsible for immigration increases. It also shows how this evidence resonated with other politicians, particularly the chart's co-option by the Trump 2024 campaign, demonstrating the agility of data visualizations to be modified to fit evolving discourse over time.

5.4.3 Remixing Screenshots of Governmental Dashboards. In our third case study, we examine a long, branching lineage. This lineage involves the same starting point: a CBP public data dashboard, where users can explore immigration data with different filters and subsets. We observe several branches from this starting point with users creating evidentiary screenshots that align with their frames and current political events. Screenshots from this dashboard appear in two key immigration-related events in January and April 2024, with users curating and annotating them from the dashboard to support different narratives, summarized in Fig 27.

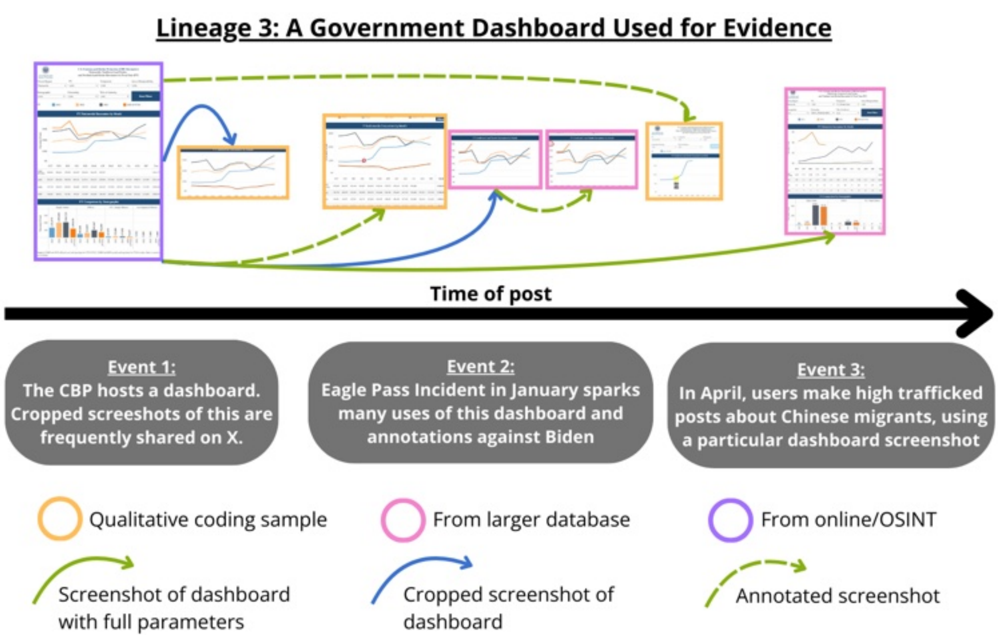

Fig. 27. A summary of two different branches from the government dashboard from current events at the border, including the Eagle Pass Incident in January and increased reports about Chinese migration in April.

Event 1: CBP hosts a dashboard and it frequently makes evidence

Thousands of these dashboard screenshots exist online and in our larger data collections, making this lineage the largest and oldest, with Fig 28 showing the CBP-hosted dashboard and an example screenshot. Unlike the two previous lineages where the creators were directly created and iterated upon their charts, this third lineage is populated by users who did not use specialized chart or illustration software. Rather, this branching lineage has an infinite number of possibilities for users to make evidence via screenshotting and remixing evidence from this government portal.

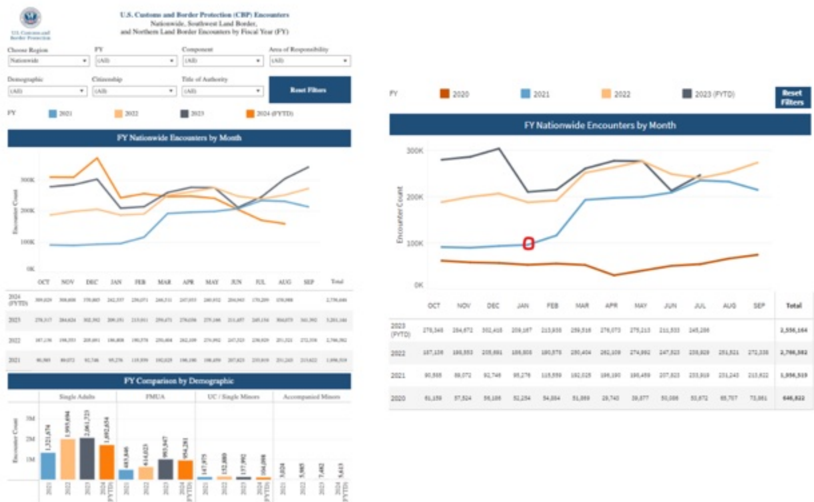

Fig. 28. CBP dashboard (left) and a cropped and annotated screenshot posted to X (right).

Mot commonly, we observe creators using cropped screenshots, supportive annotations, and the affordances of the dashboard to create charts to fit their frames and respond to different immigration-related events.

Event 2: Eagle Pass Incident prompts users to make evidence with this dashboard

In January 2024, several users created and reposted evidence generated from screenshots of this dashboard following the Eagle Pass incident (see §2). These images (Fig 29) show line chart snapshots alongside portions of the CBP dashboard, created from various inputs and filters. Tactics like cherry-picking dashboard elements, applying annotations, and cropping charts are common in these posts. Notably, many cropped screenshots omitted the data filters used, leaving out important context and risking (or inviting) misinterpretation. Annotations were often used to frame the Biden administration as culpable or to highlight peaks in immigration numbers tied to specific Southern border events, a recurring tactic in our study.

Fig. 29. Examples of January posts by users with selective cropping and annotations supporting the timeline of early January 2024 having increased migration and implying the Biden administration's culpability.

Event 3: High engagement posts about Chinese migrants

In April 2024, this dashboard was used to produce evidence spurred by reports of a large number of single male migrants from China at the US-Mexico border across major news outlets [11, 140] and from a Homeland Security Announcement [95]. Fig 30 demonstrates a popular screenshot supporting this framing of the event, where users posted longer screenshots of the dashboard to highlight migration of Chinese nationals via filtering of specific nationalities. In this case, having more of the dashboard visible allowed users to sharpen their evidence to support a subframe of criminality (discussed in §5.3.3) that a large, dangerous group of Chinese military-aged men were coming to the United States as part of a pseudo-military action.

Fig. 30. Users leveraged dashboard features to produce evidence towards a sub-framing around Chinese national migration, filtering to only show Chinese migration over time, stressing the increase between years.

This lineage demonstrates that through the use of a free online tool provided by the CBP and publicly available data, motivated individuals without extensive expertise can create data visualizations to use as supporting evidence for their preferred frames about immigration into the US. In several cases, we see visualizations created in response to real-world events that provide opportunity for both garnering attention and participating in "framing contests" [29, 52, 183].

We also note that the strategies used to frame data visualizations in this third lineage are similar to those in the previous two lineages (supportive annotations, cropping or extending details of visualizations), but with a different method of evidence generation (screenshotting from a dashboard instead of creating charts with source data).

6 Discussion

In this paper, we investigated how creators and social media users engage in loosely collaborative efforts to create and modify data visualizations to support anti-immigrant frames. Focusing our analysis upon three case studies of data visualization lineages (DVLs), which consist of a source data visualization and any subsequent modifications, we identified several strategies used to create and adapt visualizations to promote specific political frames.

Next, we explore some of the implications of our work, highlighting the value of determining and making salient the provenance of data visualizations as visual objects, the participatory nature that drives the creation and evolution of these visualizations, and discussing some of the opportunities for supporting this kind of work.

6.1 The Political and Material Offline Impacts of Data Visualization Lineages

Our findings underscore the importance of identifying and analyzing data visualization lineages within online political discourse—including unraveling the provenance of data visualizations and revealing how specific choices in their creation and modification reflect efforts to fit the data to political frames that shape how people interpret the underlying data.

In the current US political climate, data visualizations have become centerpieces within political messaging about migration. In our introduction, we called attention to the presence of our second data visualization lineage at Donald Trump's campaign rally in Butler, Pennsylvania in July 2024. As we noted earlier, this data visualization lineage had been used previously as evidence in Senate hearings by Senator Ron Johnson to support claims around the influx of migrants during the Biden administration. After the assassination attempt on former President Trump, this particular visualization grew immensely in popularity. Through our data visualization lineages, we demonstrate how the Trump campaign co-opted this set of charts as an "objective" representation of the migrant crisis, and leveraged "numbers" as evidence to promote a frame that positions the influx of migrants as the result of Biden's border policies.

We also observe how data visualizations can be created or updated in response to current events and subsequently become part of real-time political discourse. We tie this to prior CSCW work on sensemaking and the iteration of content during real time, political events and conflict [159, 180, 193]. In our third lineage, we saw this explicitly with screenshots of the CBP dashboard filtered to display migration of Chinese Nationals (Fig 29, above) after mass media coverage [11] of a "Chinese migrant surge". We recently witnessed this dynamic when the creator of the first lineage created a chart explicitly highlighting Haitian immigration after former President Trump made disparaging and false allegations about Haitian migrants in Springfield, Ohio [131]. This resulted in material threats towards Haitian residents in Springfield, Ohio [64, 85, 89]. Unfortunately, these data visualizations support rhetoric that has troubling real life impacts, such as threats inciting violence against immigrant groups [17, 20, 55, 77, 78, 99, 105, 162]. Elected political leaders have also used these visualizations to influence public policy, as in our second lineage.

With migration to the US growing as a result of forced displacement due to climate change, war, famine, and other disasters [24, 96, 129], we anticipate that anti-immigrant frames will continue to be popular and data-driven evidence to support those frames, regardless of accuracy, will proliferate. Furthermore, as surveillance of marginalized groups increases, particularly at the US-Mexico border [39, 137, 161], so will the amount of data available about these groups, which could lead to more sophisticated data visualizations, framing tactics, and DVLs emerging from them. And as AI systems (such as X's Grok) become equipped with more data visualization generation features, we anticipate an influx in the creation of visualizations used to spread harmful political rhetoric towards vulnerable groups.

Researchers in CSCW and related fields should continue to study how data visualization tools are appropriated to create and modify visualizations to support political frames, especially in cases where the visualizations are misleading and the frames harmful to marginalized people. Understanding this as collaborative work (as prior CSCW work has done [159, 178, 193]) that occurs amongst online audiences, creators, and political figures as digital publics make sense of complex, evolving issues may help to detect and migate information operations regarding such issues. We also recommend that journalists and political activists from other perspectives consider adopting the broader strategy of collaborative data visualization as political communication to demonstrate how the same underlying data can be used to support other frames, including some that counter the anti-immigration frames that dominated our analysis.

6.2 Tensions involving Open Data and Political Misinformation

The challenges do not rest solely in the tools that create data visualizations, but also in the underlying data. Our work surfaces tensions between US government requirements to have open and transparent data and the potential for this data to be visualized in ways that mislead [51, 57].

These tensions emerge from the requirement that federal agencies make their collections of demographic data available for the general public to interact with or via data visualizations on government websites [51, 57]. This demographic data is often collected through official federal data collection efforts such as the census, the CDC, and CBP. Because the US government is responsible for defining what data is collected, measurement units in these open datasets are often kept intentionally broad to serve multiple public policy purposes. Transparency is an important part of democracy and these data are central to public policy planning and decision-making in the US, including the allocation of financial resources and the distribution of congressional districts.

However, open demographic data can also be crafted into visualizations that serve as political propaganda and/or misinformation. Notably, the overwhelming majority of visualizations in our dataset rely on public data to represent migration trends from non-WEIRD countries, which are framed to fit xenophobic arguments around demographic shifts. As Lee et al. demonstrate in the context of COVID-19 data visualizations, growing public distrust in federal agencies increases the potential for open data to be re-appropriated to generate misinformation [116].

This appropriation can occur through publicly available data as well as government-provided dashboards that allow members of the public to visualize and interact with the data, and easily create screenshots that can be marked up to fit political frames, as seen in our third lineage. Creators have also started to leverage the capabilities of AI to generate visualizations of this open data. A pertinent example in this study is of our first lineage's creator making a chart about immigration (Fig 30) and claiming Grok, the X platform's AI, is revealing hidden governmental data. As AI can make mistakes and hallucinate [37, 153, 194], AI generated outputs interpreting or citing governmental data may exacerbate this tension of open data and potential misinformation. We discuss implications for research and design to better understand and address the misappropriation of open data below.

6.3 Annotations and Remixes to Data Visualizations to Promote Strategic Political Frames and Mislead

Our work departs from and extends existing research on misleading data visualizations in important ways. Due to the implicit assumptions of objectivity associated with data and numbers, visualizations are often considered to be neutral. However our work, and the work of scholars such as Lee et al. [116] and Desrosières [66] show that visualizations often reflect the intention of their creators [60, 69, 76]. This intent can be to persuade or mislead their audiences.

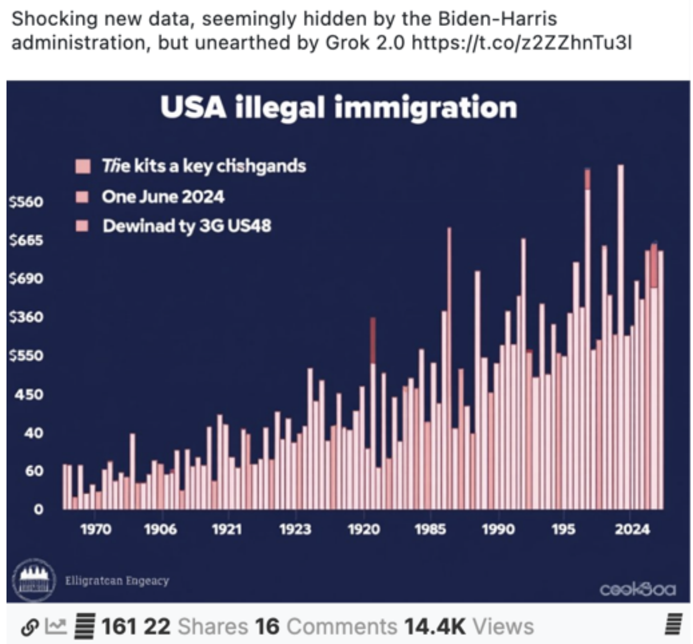

Fig. 31. An example chart from a now deleted post made from Grok 2.0 and posted on X by the creator we covered in our first data visualization lineage, retrieved from Junkipedia.

Lisnic et al. [122] offer a typology of "misleading" visualizations and focus on the technical choices made to existing charts that are vulnerable to being manipulated. In particular, Lisnic et al. [121, 122] discuss reasoning errors such as cherry-picking, setting an arbitrary threshold, and making a misleading causal inference. Additionally, they discuss visual and textual annotations and screenshotting as construction attributes of misleading visualizations.

Our work extends this understanding of misleading visualizations through DVLs, which examine these tactics within an iterative and loosely collaborative process, revealing the order of specific decisions made by creators and annotators to not just mislead but support or counter certain (anti-immigrant) political frames. We demonstrate how political annotations, modifications, and iterations to data visualizations are repeatedly used to draw attention to current events, and to produce/adapt evidence to support existing political frames, even as new events unfold and new data is produced.

Our findings on the use of annotations and remixes align with Lee et al.'s [116] findings on how social media users co-opt common visualization strategies to propose a counter-narrative to mainstream reporting of the impact of COVID-19. In our case, creators in our data visualization lineages co-opted common annotation practices alongside additional tactics such as erasing parts of the source data visualization via photo editing, treating data visualizations as iterative visual objects, extending the annotational grammar in previous work. As Lee et al. [116] show in the case of COVID-19, social media users created counter data visualizations that help legitimize alternative claims. Likewise, in our work, social media creators shared visualizations that supported anti-immigrant frames, often through strategic presentation of the underlying data that brought the data into conversation with strategically activated concerns about criminality, welfare funding, shifting demographics, and an allegedly incompetent Presidential administration.

Though this paper focuses on one specific discourse related to immigration, similar tactics can and are being used in the creation and adaptation of data visualizations to support other politically salient frames [118, 197]. Tools for creating these visualizations, or dashboards where existing visualizations can become screenshots, are readily available and audiences appear to be primed to view and use them to support their preferred frames. Future work should explore their use within discussions of other politically fraught topics, including political discourse around gender-affirming and reproductive health care in the US.

7 Implications for Future Research and Design

Based on our investigation, we propose a number of research opportunities to support researchers in identifying data visualization lineages and mitigating their impact in online discourse.

7.1 Tracing Provenance

Provenance analysis involves understanding the different steps that took place in transforming a specific media from its original state to its current state [107, 116, 145]. Our work combines long-established methods, such as OSINT frameworks and qualitative analysis, to perform provenance analysis on data visualization lineages (DVLs). We believe provenance and the understanding of DVLs can greatly benefit the study of data visualizations, and other political media. We also believe that data visualizations, which are visual objects that come to be from multiple points of origin (the data powering them, the visualization creator, and visualization annotations) are a rich space for provenance research.

Built-in provenance standards are gaining more research and traction in the growing landscape of AI-generated visual media [192], as individuals can more easily create more content with veneers of legitimacy and credibility, such as data visualizations. The Coalition for Content Provenance and Authenticity (C2PA) and their C2PA standard of Content Credentials (CC), which experts often point to as the most promising standard [39, 64], are in the early stages of being adopted by large companies such as Adobe, Google, and OpenAI. Although these standards exist, we do not know how and to what extent companies adopt them in their AI-supported products. Additionally, even with this adoption, provenance is multidimensional and not always well understood by users, necessitating careful development and implementation in online spaces [84].

We join prior work in CSCW such as Feng et al. and Zade et al. [84, 196] in advocating for more research interventions around the development and adoption of provenance at the point of media creation and post-creation iterations to support media literacy and transparency. We argue for interventions in user platforms and independent analysis tools to support researchers, and the public, in detecting the provenance of data visualizations and other media objects.

Particularly, we see opportunities to develop tools and techniques for tracing the provenance of open data visualizations from official governmental sources in the cases of their spread during crises or current events in potentially misleading or even harmful ways. For example, government agencies may consider enabling screenshot watermarks in both visual and metadata levels on their dashboards to help attribute evidence created from them. These dashboards and agencies may also consider adding watermarks or educational resources during key crisis moments or current events, such as the highly covered Eagle Pass incident in January 2024 [58, 63], since social media users often try to engage in sensemaking and evidence production during acute times of crisis, as seen in our and past CSCW work [159, 160, 178, 193]. We also advocate for governmental dashboards to have provenance measures to avoid nefarious misappropriation of this open data.

Additionally, improvements to existing OSINT frameworks and methods could be fruitful in navigating this new space. Reverse image search engines such as Google Lens may consider these types of remixing and provenance notions to better aid OSINT efforts and public understanding

Fig. 32. A Google Lens output in one of our lineages, which does not prioritize chronological order or notions of provenance such as editing.

of how visual information changes by returning results chronologically and by clusters of spread. These engines may also consider providing context for data visualizations and images that replicate lineage behavior, such as moving between social media platforms or being slightly modified, as in our second lineage. Prioritizing image provenance and movement across networks in such information seeking interfaces would help the OSINT research community and the broader public understand these information dynamics.

We also see opportunities for future work to consider how audiences perceive and interpret AI-generated visualizations, particularly of data from official US government sources. Future work can investigate the impacts of using AI systems to create visualizations of open federal datasets or if AI systems can include built-in safeguards against these kinds of visualizations. Additionally, considering ongoing legal issues around data ownership [163], the question remains about who should be held accountable for anti-immigrant content generated by AI systems once these artifacts enter online spaces and may again be remixed, reframed, and amplified for political gain.

7.2 Defining Harms

Our work identifies data visualization lineages that contribute to problematic political discourse, often using open data from federal sources. We view these anti-immigrant visualizations as having the potential to cause physical and material harm to the groups being degraded by this rhetoric, as other researchers have found [77, 78, 99], and with concerning increases in targeted crimes against Haitian migrants in Ohio [64, 85, 89] as well as recent raids on voting organizations in Hispanic communities [187]. Immigrant communities are vulnerable not only to such hate crimes, but also to unjust and cruel deportation incidents [46, 162].

However, we note that it has been difficult throughout this paper to specify and call out specific harms related to misleading visualizations, in part due to a lack of shared definitions across the different fields pulled together by this research.

We therefore note research opportunities to investigate and characterize the various harms that may result from data visualization lineages used within political rhetoric online. We look to work done by Dev et al. [67] in their framework codifying and understanding the types of harms in AI models, including stereotyping, erasure, quality of service, dehumanization and disparagement—which can serve as a starting point for research into the various kinds of harms resulting from data visualizations that support dehumanizing political frames. It is also imperative to understand how such data visualizations may fit into schemas of online harm types and alongside other content towards potential moderation [102, 168].

Care also needs to be taken to center the groups subject to political frames, such as anti-immigrant rhetoric, when defining the harms to which they may be exposed. Alongside further research into the creators, remixers, and amplifiers of these kinds of images, we believe expanding the purview of research in our fields to focus on documenting, uplifting, and repairing the harms experienced from these phenomena is needed (echoing Schafer et al. [166]). Additionally, since data visualizations often claim credibility and objectivity, we call for more research into (1) content moderation and community guidelines on social media platforms related to partisan creators and their use of open governmental data to levy misleading political claims against marginalized groups, and (2) exploring how data visualizations may be subject to community moderation guidelines as other visual objects.

8 Limitations and Future Work

We acknowledge several limitations in this study. Our data itself has limitations, from our purposive sampling, to the original data being sourced from Junkipedia, which does not have broad data coverage like the X API—the fees for which were far beyond our research budget. Additionally, our computationally-assisted methods also introduced margins of error, such as our classification process's accuracy of 88%. This work also only included charts in English, despite our larger Junkipedia collection containing English and Spanish data.

We also acknowledge that while we posit DVLs as an important methodological approach, collecting and organizing the lineages in this paper required a large amount of manual work to assemble the lineages, which could be further automated with additional computational assistance.

We focus on the general dynamics and tactics of data visualization design and remixing towards political online discourse, but we acknowledge this data is anchored in one domain of study — anti-immigrant rhetoric — and in a particular composition of data visualizations that favored line and bar charts. This study is also focused on X data, a choice made in other studies for its salience on political conversation [159] and here for how Musk, the network's owner, has boosted this content on such a large network, as seen in this work.

We hope this study can inform and inspire future work investigating how data visualizations can become propaganda in other issues, platforms, and even types of visualizations. We also hope that our treatment of data visualizations as iterative, visual objects inspires new work in data visualization, visual culture, social media, and decision making. Lastly, as open governmental datasets are rapidly being removed under the Trump administration in 2025 [134], future work can investigate the impacts of this removal of open datasets on visual propaganda.

9 Conclusion

In this paper, we explored how data visualizations were used within social media discourse, often in misleading ways, to support anti-immigrant frames around the 2024 US-Mexico border crisis. We introduce the concept of data visualization lineages (DVLs) to highlight how these visualizations were iterated upon, adapted, remixed, and spread by content creators, political operatives, and online influencers. Our work demonstrates that even the most "neutral" of data visualizations are not apolitical and that data and data visualizations can be appropriated to support nefarious political frames that mislead and promote harmful propaganda. Furthermore, we demonstrate the participatory and visual nature of data visualizations as they are iterated upon and spread online.

Noting the role of government-provided data in the generation of these visualizations, and the frequent disconnect between the original data and downstream visualizations, we provide recommendations around identifying and communicating provenance as a way to mitigate misinterpretation and mischaracterization within data visualization lineages. We hope this work will spark further conversations about the role of data visualizations in supporting harmful and misleading propaganda targeting marginalized groups and inform research and design to mitigate these harms.

Acknowledgments

References

[1] [n. d.]. Coalition for Content Provenance and Authenticity (C2PA). https://c2pa.org/. Accessed: 2024-10-27.

[2] [n. d.]. The Great Replacement - Global Project Against Hate and Extremism. https://globalextremism.org/the-great-replacement

[3] 2023. Title 42 has ended. Here's what it did, and how US immigration policy is changing. https://apnews.com/article/immigration-biden-border-title-42-mexico-asylum-be4e0b15b27adb9bede87b9bbefb798d

[4] 2024. CBP Public Data Portal | U.S. Customs and Border Protection. https://www.cbp.gov/newsroom/stats/cbp-public-data-portal

[5] 2024. FEMA's Role in Migrant Assistance: Exploring the Shelter and Services Program. https://www.americanimmigrationcouncil.org/research/shelter-services-program

[6] 2024. Glossary | OHSS - Office of Homeland Security Statistics. https://ohss.dhs.gov/glossary

[7] 2024. ICYMI: Sen. Johnson Fights Disastrous Ukraine Funding Bill That Does Not Secure the U.S. Border. https://www.ronjohnson.senate.gov/2024/2/icymi-sen-johnson-fights-disastrous-ukraine-funding-bill-that-does-not-secure-the-u-s-border Section: Press Releases.

[8] 2024. ICYMI: Sen. Johnson Presents Evidence of Sec. Mayorkas's Crimes. https://www.ronjohnson.senate.gov/2024/4/icymi-sen-johnson-presents-evidence-of-sec-mayorkas-s-crimes Section: Press Releases.

[9] 2024. OSINT Framework. https://osintframework.com/

[10] 2024. State and Local Expenditures | Urban Institute. (2024). https://www.urban.org/policy-centers/cross-center-initiatives/state-and-local-finance-initiative/state-and-local-backgrounders/state-and-local-expenditures

[11] 2024. Surge of Chinese migrants crossing Southern border into the U.S. https://www.nbcnews.com/nightly-news/video/surge-of-chinese-migrants-crossing-southern-border-into-the-u-s-209500741912

[12] 2024. Welfare Benefits for Non-citizens. https://www.brookings.edu/articles/welfare-benefits-for-non-citizens/

[13] 2024. Welfare benefits or Temporary Assistance for Needy Families (TANF) | USAGov. https://www.usa.gov/welfare-benefits

[14] Stuart Altman and Marion Ein Lewin. 2000. America's health care safety net: intact but endangered. (2000).

[15] Glenn C. Altschuler and Stuart M. Blumin. 2022. Perspective | The 'great replacement' theory rises again, ending in tragedy. *Washington Post* (June 2022). https://www.washingtonpost.com/outlook/2022/06/02/great-replacement-theory-rises-again-ending-tragedy/

[16] Marcos Montoya Andrade and Dr. Ray Serrano. 2024. A New Wave of Hate. https://lulac.org/a_new_wave_of_hate/

[17] Carlos Arcila Calderón, Patricia Sánchez Holgado, Jesús Gómez, Marcos Barbosa, Haodong Qi, Alberto Matilla, Pilar Amado, Alejandro Guzmán, Daniel López-Matías, and Tomás Fernández-Villazala. 2024. From online hate speech to offline hate crime: the role of inflammatory language in forecasting violence against migrant and LGBT communities. *Humanities and Social Sciences Communications* 11, 1 (Oct. 2024), 1–14. https://doi.org/10.1057/s41599-024-03899-1 Publisher: Palgrave.

[18] Nell Salzman Armando Sanchez. 2024. Experts say anti-immigrant rhetoric led to viral — and incorrect — allegations of a migrant takeover. https://www.chicagotribune.com/2024/09/03/experts-say-anti-immigrant-rhetoric-led-to-viral-and-incorrect-allegations-of-a-migrant-takeover/

[19] Isaac Arnsdorf. 2024. Trump baselessly blames assassination attempt on Democratic messaging. *Washington Post* (Aug. 2024). https://www.washingtonpost.com/politics/2024/08/03/trump-shooting-democrats-democracy/

[20] Emelynn Arroyave and Oprah Cunningham. 2025. The Human Costs of Trump's Immigration Crackdown. *The Leadership Conference on Civil and Human Rights* (2025). https://civilrights.org/blog/the-human-costs-of-trumps-immigration-crackdown/

[21] Kwasi Asiedu. 2024. Tucker Carlson is wrong when comparing migrants illegally entering the US and US births. https://www.poynter.org/fact-checking/2024/invasion-illegal-immigrants-outpace-american-births/

[22] Ruben T. Azevedo, Sophie De Beukelaer, Isla L. Jones, Lou Safra, and Manos Tsakiris. 2021. When the lens is too wide: The political consequences of the visual dehumanization of refugees. *Humanities and Social Sciences Communications* 8, 1 (May 2021), 1–16. https://doi.org/10.1057/s41599-021-00786-x Number: 1 Publisher: Palgrave.

[23] Illya Bakurov, Marco Buzzelli, Raimondo Schettini, Mauro Castelli, and Leonardo Vanneschi. 2022. Structural similarity index (SSIM) revisited: A data-driven approach. *Expert Systems with Applications* 189 (2022), 116087.

[24] Andrew Baldwin. 2017. Climate change, migration, and the crisis of humanism. *WIREs Climate Change* 8, 3 (2017), e460. https://doi.org/10.1002/wcc.460 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.460.

[25] Christine A Barry, Nicky Britten, Nick Barber, Colin Bradley, and Fiona Stevenson. 1999. Using reflexivity to optimize teamwork in qualitative research. *Qualitative health research* 9, 1 (1999), 26–44.

[26] Marco Bastos, Dan Mercea, and Fabio Goveia. 2023. Guy next door and implausibly attractive young women: The visual frames of social media propaganda. *New Media & Society* 25, 8 (2023), 2014–2033.

[27] Jean-Louis Baudry and Alan Williams. 1974. Ideological effects of the basic cinematographic apparatus. *Film Quarterly* 28, 2 (1974), 39–47.

[28] Brian Bell, Francesco Fasani, and Stephen Machin. 2013. Crime and Immigration: Evidence from Large Immigrant Waves. *The Review of Economics and Statistics* 95, 4 (Oct. 2013), 1278–1290. https://doi.org/10.1162/REST_a_00337

[29] Robert D Benford. 2013. Frame disputes. *The Wiley-Blackwell Encyclopedia of Social and Political Movements* (2013), 1–3.

[30] Robert D Benford and David A Snow. 2000. Framing processes and social movements: An overview and assessment. *Annual review of sociology* 26, 1 (2000), 611–639.

[31] Ken Bensinger. 2024. A Mysterious Flier, a Tiny Charity and a Disinformation Campaign at the Border - The New York Times. https://www.nytimes.com/2024/04/30/us/politics/immigration-disinformation-campaign-biden-trump.html

[32] Carl T Bergstrom and Jevin D West. 2021. *Calling bullshit: The art of skepticism in a data-driven world.* Random House Trade Paperbacks.

[33] Charles Berret and Tamara Munzner. 2024. Iceberg Sensemaking: A Process Model for Critical Data Analysis. *IEEE Transactions on Visualization and Computer Graphics* (2024), 1–18. https://doi.org/10.1109/TVCG.2024.3486613

[34] Kratika Bhagtani, Amit Kumar Singh Yadav, Emily R. Bartusiak, Ziyue Xiang, Ruiting Shao, Sriram Baireddy, and Edward J. Delp. 2022. An Overview of Recent Work in Media Forensics: Methods and Threats. arXiv:2204.12067 [cs.CV] https://arxiv.org/abs/2204.12067

[35] Aparna Bharati, Daniel Moreira, Joel Brogan, Patricia Hale, Kevin Bowyer, Patrick Flynn, Anderson Rocha, and Walter Scheirer. 2019. Beyond Pixels: Image Provenance Analysis Leveraging Metadata. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 1692–1702. https://doi.org/10.1109/WACV.2019.00185

[36] Melissa Bica, Leysia Palen, and Chris Bopp. 2017. Visual Representations of Disaster. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1262–1276. https://doi.org/10.1145/2998181.2998212

[37] Abeba Birhane, Vinay Uday Prabhu, and Emmanuel Kahembwe. 2021. Multimodal Datasets: Misogyny, Pornography, and Malignant Stereotypes. *arXiv e-prints* (2021), arXiv–2110. https://arxiv.org/abs/2110.01963

[38] Roland Bleiker, David Campbell, Emma Hutchison, and Xzarina Nicholson. 2013. The visual dehumanisation of refugees. *Australian Journal of Political Science* 48, 4 (Dec. 2013), 398–416. https://doi.org/10.1080/10361146.2013.840769 Publisher: Routledge _eprint: https://doi.org/10.1080/10361146.2013.840769.

[39] Geoffrey A Boyce. 2016. The rugged border: Surveillance, policing and the dynamic materiality of the US/Mexico frontier. *Environment and Planning D: Society and Space* 34, 2 (2016), 245–262.

[40] Sarah Bracke and Luis Manuel Hernández Aguilar. 2024. The politics of replacement: from "race suicide" to the "great replacement". In *The Politics of Replacement.* Routledge, 1–19.

[41] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[42] Virginia Braun and Victoria Clarke. 2023. Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a *knowing* researcher. *International Journal of Transgender Health* 24, 1 (Jan. 2023), 1–6. https://doi.org/10.1080/26895269.2022.2129597

[43] Philip Bump. 2024. The political exploitation of immigration data. https://www.washingtonpost.com/politics/2024/01/04/immigration-republicans-speaker-johnson/

[44] Philip Bump. 2024. Updating (and fixing) Trump's favorite chart. *The Washington Post* (Oct. 2024). https://www.washingtonpost.com/politics/2024/10/24/trump-biden-immigration/ Accessed: 2024-10-27.

[45] Desiree Butler. 2024. Trump finds familiar fault lines over immigration in presidential campaign rhetoric. *Associated Press News* (2024). https://apnews.com/article/immigration-trump-chart-butler-border-67866bd05a894a6bea0d670de2adc4f7 Accessed: 2024-10-27.

[46] A Policy Statement by the Society for Community Research and Action: Division 27 of the American Psychological Association. 2018. Statement on the effects of deportation and forced separation on immigrants, their families, and communities. *American Journal of Community Psychology* 62, 1-2 (2018), 3–12.

[47] Alberto Cairo. 2019. *How Charts Lie: Getting Smarter about Visual Information.* W. W. Norton & Company. Google-Books-ID: qP2KDwAAQBAJ.

[48] Céline Cantat, Antoine Pécoud, and Hélène Thiollet. 2023. Migration as Crisis. *American Behavioral Scientist* (July 2023), 00027642231182889. https://doi.org/10.1177/00027642231182889 Publisher: SAGE Publications Inc.

[49] S.K. Card and J. Mackinlay. 1997. The structure of the information visualization design space. In *Proceedings of VIZ '97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium.* IEEE Comput. Soc, Phoenix, AZ, USA, 92–99,. https://doi.org/10.1109/INFVIS.1997.636792

[50] Stephen Castles. 2013. The Forces Driving Global Migration. *Journal of Intercultural Studies* 34, 2 (April 2013), 122–140. https://doi.org/10.1080/07256868.2013.781916 Publisher: Routledge _eprint: https://doi.org/10.1080/07256868.2013.781916.

[51] Census Bureau. 1997. About the Topic of Race. Retrieved September 9, 2022 from https://www.census.gov/topics/population/race/about.html

[52] Kevin M. Carragee Charlotte Ryan and William Meinhofer. 2001. Theory Into Practice: Framing, the News Media, and Collective Action. *Journal of Broadcasting & Electronic Media* 45, 1 (2001), 175–182. https://doi.org/10.1207/s15506878jobem4501_11

[53] Kathy Charmaz. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. sage.

[54] Kathy Charmaz and Robert Thornberg. 2021. The pursuit of quality in grounded theory. *Qualitative research in psychology* 18, 3 (2021), 305–327.

[55] Ko Lyn Cheang. 2025. ACLU sues Border Patrol over California immigration raid that set off panic among farmworkers. *San Francisco Chronicle* (2025). https://www.sfchronicle.com/california/article/aclu-border-patrol-lawsuit-20190436.php

[56] Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20, 1 (1960), 37–46.

[57] United States Equal Employment Opportunity Commission. 1966. EEO-1 Component 1 Data Collection. Retrieved September 9, 2022 from https://eeocdata.org/eeo1

[58] Stephanie Corpi. 2024. Drowning deaths of several migrants at US-Mexico border heightens tensions even more. *NPR* (Jan. 2024). https://www.npr.org/2024/01/18/1225446403/drowning-deaths-of-several-migrants-at-us-mexico-border-heightens-tensions-even-

[59] Annie Correal. 2020. As Pandemic Reveals Gaps in Safety Nets, 'You Can't Look Away'. *The New York Times* (2020). https://www.nytimes.com/2020/12/03/neediest-cases/as-pandemic-reveals-gaps-in-safety-nets-you-cant-look-away.html Accessed: 2024-10-27.

[60] Michael Correll. 2019. Ethical Dimensions of Visualization Research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300418

[61] Dharma Dailey and Kate Starbird. 2015. "It's Raining Dispersants": Collective Sensemaking of Complex Information in Crisis Contexts. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW'15 Companion)*. Association for Computing Machinery, New York, NY, USA, 155–158. https://doi.org/10.1145/2685553.2698995

[62] Viorela Dan, Britt Paris, Joan Donovan, Michael Hameleers, Jon Roozenbeek, Sander van der Linden, and Christian von Sikorski. 2021. Visual mis-and disinformation, social media, and democracy. *Journalism & Mass Communication Quarterly* 98, 3 (2021), 641–664.

[63] David Martin Davies. 2024. This Texas city is caught in a fight over who controls the southern border. *NPR* (Feb. 2024). https://www.npr.org/2024/02/07/1229856703/this-texas-city-is-caught-in-a-fight-over-who-controls-the-southern-border

[64] Sydney Dawes. 2024. Hate groups converge on Springfield after false claims about Haitian immigrants. https://www.daytondailynews.com/local/hate-groups-converge-on-springfield-following-false-claims-about-haitian-immigrants/R6BSUOAMGZDNZNSV4PUTBDPOSY/

[65] Daniel Dayan. 1974. The tutor-code of classical cinema. *Film Quarterly* 28, 1 (1974), 22–31.

[66] Alain Desrosières. 1998. *The politics of large numbers: A history of statistical reasoning*. Harvard University Press.

[67] Sunipa Dev, Emily Sheng, Jieyu Zhao, Aubrie Amstutz, Jiao Sun, Yu Hou, Mattie Sanseverino, Jiin Kim, Akihiro Nishi, Nanyun Peng, and Kai-Wei Chang. 2022. On Measures of Biases and Harms in NLP. http://arxiv.org/abs/2108.03362 arXiv:2108.03362 [cs].

[68] Priya Dhawka and Sayamindu Dasgupta. 2025. The Social Construction of Visualizations: Practitioner Challenges and Experiences of Visualizing Race and Gender Demographic Data. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–16.

[69] Catherine D'Ignazio and Lauren F Klein. 2020. 1. The Power Chapter. In *Data Feminism*. The MIT Press, Cambridge, Massachusetts.

[70] Evanthia Dimara, Steven Franconeri, Catherine Plaisant, Anastasia Bezerianos, and Pierre Dragicevic. 2020. A Task-Based Taxonomy of Cognitive Biases for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 26, 2 (2020), 1413–1432. https://doi.org/10.1109/TVCG.2018.2872577

[71] Priya Dixit. 2022. Red Pills, White Genocide, and "the Great Replacement": Rewriting History, and Constructing White Victimhood in/through Far-Right Extremist Manifestos and Texts. In *Race, Popular Culture, and Far-right Extremism in the United States*, Priya Dixit (Ed.). Springer International Publishing, Cham, 173–205. https://doi.org/10.1007/978-3-031-10820-4_6

[72] Hulya Dogan, Kiet Nguyen, and Ismini Lourentzou. 2024. Narrative Characteristics in Refugee Discourse: An Analysis of American Public Opinion on the Afghan Refugee Crisis After the Taliban Takeover. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1 (April 2024), 212:1–212:31. https://doi.org/10.1145/3653703

[73] Hannah Dormido, Hannah Knowles, Rep. Marianna Sotomayor, and Jacob Bogage. 2024. A look at Trump's misleading and inaccurate graph on U.S. immigration. https://www.washingtonpost.com/politics/2024/05/23/look-trumps-misleading-inaccurate-graph-us-immigration/ Accessed: 2024-10-27.

[74] Nicholas Dufour, Arkanath Pathak, Pouya Samangouei, Nikki Hariri, Shashi Deshetti, Andrew Dudfield, Christopher Guess, Pablo Hernández Escayola, Bobby Tran, Mevan Babakar, et al. 2024. AMMeBa: A Large-Scale Survey and Dataset of Media-Based Misinformation In-The-Wild. *arXiv preprint arXiv:2405.11697* (2024).

[75] Richard Dyer. 2017. The light of the world. In *White* (2nd ed.). Routledge, New York.

[76] Marian Dörk, Patrick Feng, Christopher Collins, and Sheelagh Carpendale. 2013. Critical InfoVis: exploring the politics of visualization. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. ACM, Paris France, 2189–2198. https://doi.org/10.1145/2468356.2468739

[77] Maureen A. Eger and Susan Olzak. 2023. The Polarizing Effect of Anti-Immigrant Violence on Radical Right Sympathies in Germany. *International Migration Review* 57, 2 (June 2023), 746–777. https://doi.org/10.1177/01979183221126461 Publisher: SAGE Publications Inc.

[78] Mattias Ekman. 2019. Anti-Immigrant Sentiments and Mobilization on the Internet. Sage Publications, 551–562. https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-195348

[79] Sarah Ellison, Amy Gardner, and Clara Ence Morse. 2024. Elon Musk's misleading election claims reach millions and alarm election officials. *Washington Post* (Sept. 2024). https://www.washingtonpost.com/politics/2024/09/10/elon-musk-illegal-immigrant-voting-misinformation/

[80] Ximena Bustillo Emma Bowman, Joe Hernandez and Jeongyoon Han. 2024. Biden calls for unity after attempted assassination of Trump : NPR. https://www.npr.org/2024/07/14/g-s1-10305/trump-assassination-attempt-biden-unity

[81] Robert M. Entman. 1993. Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication* 43, 4 (Dec. 1993), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

[82] Hany Farid. 2022. Creating, Using, Misusing, and Detecting Deep Fakes. *Journal of Online Trust and Safety* 1, 4 (Sep. 2022). https://doi.org/10.54501/jots.v1i4.56

[83] Hany Farid. 2024. How Princess Kate and Photo Forensics Could Usher In a New Era of Digital Provenance. *TIME* (2024). https://time.com/6899993/princess-kate-middleton-photo-forensics-digital-provenance-credentials/

[84] K. J. Kevin Feng, Nick Ritchie, Pia Blumenthal, Andy Parsons, and Amy X. Zhang. 2023. Examining the Impact of Provenance-Enabled Media on Trust and Accuracy Perceptions. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 270 (Oct. 2023), 42 pages. https://doi.org/10.1145/3610061

[85] Ashleigh Fields. 2024. Springfield bomb threat used 'hateful' language towards migrants, Haitians. https://thehill.com/homenews/campaign/4877602-ohio-bomb-threat-immigrant-hate Accessed: 2024-10-27.

[86] Jasmine Garsd. 2024. Anguish and fear in Florida amid rising anti-immigrant sentiment. *NPR* (April 2024). https://www.npr.org/2024/04/24/1236631178/florida-immigration-desantis-sb1718-fear

[87] Clifford Geertz. 2017. *The interpretation of cultures*. Basic books.

[88] Julia Gelatt and Muzaffar Chishti. 2022. What Has Been the Impact of the Pandemic on Immigrant Communities and U.S. Immigration Trends and Policies Two Years On? *Migration Policy Institute* (2022). https://www.migrationpolicy.org/news/pandemic-impact-immigrants-us-immigration-two-years-on

[89] Erin Glynn. 2024. Haitian Springfield residents say they don't feel safe. https://www.dispatch.com/story/news/2024/09/14/haitian-springfield-residents-say-they-dont-feel-safe/75228724007/

[90] Judith Goetz. 2021. 'The Great Replacement': Reproduction and population policies of the far right, taking the Identitarians as an example. *DiGeSt-Journal of Diversity and Gender Studies* 8, 1 (2021).

[91] Erving Goffman. 1974. *Frame analysis: An essay on the organization of experience*. Harvard University Press, Cambridge, MA, US. Pages: ix, 586.

[92] Robert Grant. 2018. *Data visualization: Charts, maps, and interactive graphics*. Chapman and Hall/CRC.

[93] Sam Gregory. 2022. Deepfakes, misinformation and disinformation and authenticity infrastructure responses: Impacts on frontline witnessing, distant witnessing, and civic journalism. *Journalism* 23, 3 (2022), 708–729.

[94] Eli Holder and Cindy Xiong Bearfield. 2024. Polarizing Political Polls: How Visualization Design Choices Can Shape Public Opinion and Increase Political Polarization. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 1446–1456. https://doi.org/10.1109/TVCG.2023.3326512

[95] House Committee on Homeland Security. 2024. Startling Stats Factsheet: Encounters of Chinese Nationals Surpass All Fiscal Year 2023 at the Southwest Border. https://homeland.house.gov/2024/04/18/startling-stats-factsheet-encounters-of-chinese-nationals-surpass-all-fiscal-year-2023-at-the-southwest-border/. Accessed: 2024-10-27.

[96] Lawrence Huang. 2023. Climate Migration 101: An Explainer. https://www.migrationpolicy.org/article/climate-migration-101-explainer

[97] Amanda Lee Hughes and Leysia Palen. 2009. Twitter adoption and use in mass convergence and emergency events. *International journal of emergency management* 6, 3-4 (2009), 248–260.

[98] Jacob Hunter. 2023. Image Analysis and Verification. https://www.osintcombine.com/post/image-analysis-and-verification
[99] Institute for Strategic Dialogue. 2024. Quantifying extremism: A data-driven analysis of UK riot-related far-right Telegram networks. https://www.isdglobal.org/digital_dispatches/quantifying-extremism-a-data-driven-analysis-of-riot-related-far-right-telegram-networks/
[100] International Organization for Migration. 2024. World Migration Report 2024. https://worldmigrationreport.iom.int/what-we-do/foreword/foreword Accessed: 2024-10-27.
[101] Adam Isacson and Ana Lucia Verduzco. 2023. Tijuana's migrant shelters under increased attack, as the U.S. sends hundreds of migrants per day - WOLA. https://www.wola.org/analysis/tijuanas-migrant-shelters-under-increased-attack-as-the-u-s-sends-hundreds-of-migrants-per-day/
[102] Jialun Aaron Jiang, Peipei Nie, Jed R. Brubaker, and Casey Fiesler. 2023. A Trade-off-centered Framework of Content Moderation. *ACM Trans. Comput.-Hum. Interact.* 30, 1, Article 3 (March 2023), 34 pages. https://doi.org/10.1145/3534929
[103] Huo Jingnan. 2024. Conspiracy theories surge following the assassination attempt on Trump. *NPR* (July 2024). https://www.npr.org/2024/07/14/nx-s1-5039234/examining-how-narratives-about-the-trump-assassination-attempt-are-resonating
[104] Miriam Jordan. 2024. Faith-Based Groups That Assist Migrants Become Targets of Extremists. *The New York Times* (June 2024). https://www.nytimes.com/2024/06/02/us/migrants-charities-shelters-threats.html
[105] Miriam Jordan. 2024. Far From Ohio, Haitian Americans Feel the Sting of Threats in Springfield. *The New York Times* (Sept. 2024). https://www.nytimes.com/2024/09/18/us/haitian-americans-springfield-ohio-reaction.html
[106] Helen Kennedy, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. 2016. The work that visualisation conventions do. *Information, Communication & Society* 19, 6 (2016), 715–735.
[107] Christy Khoury, Alexander Owen-Smith, Una Joh, Yiran Duan, and Jeff Hemsley. 2023. Multi-Modal Crisis Discourse and Collective Sensemaking on TikTok. *Proceedings of the Association for Information Science and Technology* 60, 1 (2023), 203–212.
[108] Maya King. 2024. What's a 'Black Job'? Trump's Anti-Immigration Remarks Are Met With Derision. *The New York Times* (2024). https://www.nytimes.com/2024/06/28/us/politics/black-job-trump-immigration.html Accessed: 2024-10-28.
[109] Miles Klee. 2024. Elon Musk Doubles Down on Great Replacement Theory. https://www.rollingstone.com/culture/culture-news/elon-musk-great-replacement-conspiracy-theory-1234941337/
[110] Gary Klein, Jennifer K. Phillips, Erica L. Rall, and Deborah A. Peluso. 2007. A data-frame theory of sensemaking. In *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, 113–155. https://doi.org/10.4324/9780203810088
[111] Audrey M Kleinsasser. 2000. Researchers, reflexivity, and good data: Writing to unlearn. *Theory into practice* 39, 3 (2000), 155–162.
[112] Faye Kollig and Casey Fiesler. 2023. "Headlines rarely soothe nerves": An Analysis of News Coverage of Social Media Mental Health Research. In *Computer Supported Cooperative Work and Social Computing*. ACM, Minneapolis MN USA, 112–116. https://doi.org/10.1145/3584931.3607012
[113] Doris Kosminsky, Jagoda Walny, Jo Vermeulen, Søren Knudsen, Wesley Willett, and Sheelagh Carpendale. 2022. Belief at first sight: Data visualization and the rationalization of seeing. *Information Design Journal* (July 2022), 43–55. https://doi.org/10.1075/idj.25.1.04kos
[114] Yubo Kou, Xinning Gui, Yunan Chen, and Kathleen Pine. 2017. Conspiracy Talk on Social Media: Collective Sensemaking during a Public Health Crisis. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 61 (Dec. 2017), 21 pages. https://doi.org/10.1145/3134696
[115] P. M. Krafft and Joan Donovan. 2020. Disinformation by Design: The Use of Evidence Collages and Platform Filtering in a Media Manipulation Campaign. *Political Communication* 37, 2 (2020), 194–214. https://doi.org/10.1080/10584609.2019.1686094 arXiv:https://doi.org/10.1080/10584609.2019.1686094
[116] Crystal Lee, Tanya Yang, Gabrielle D Inchoco, Graham M. Jones, and Arvind Satyanarayan. 2021. Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3411764.3445211
[117] Tristan Lee and Kolina Koltai. 2024. OSHIT: Seven Deadly Sins of Bad Open Source Research. *Bellingcat* (April 2024). https://www.bellingcat.com/resources/2024/04/25/oshit-seven-deadly-sins-of-bad-open-source-research/ Accessed: 2024-10-27.
[118] Steven Lee Myers and Stuart A. Thompson. 2025. Falsehoods Fuel the Right-Wing Crusade Against U.S.A.I.D. *The New York Times* (2025). https://www.nytimes.com/2025/02/07/business/usaid-conspiracy-theories-disinformation.html Accessed: 2025-03-01.

[119] Elsie Lee-Robbins and Eytan Adar. 2022. Affective Learning Objectives for Communicative Visualizations. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–11. https://doi.org/10.1109/TVCG.2022.3209500

[120] Calvin A. Liang, Sean A. Munson, and Julie A. Kientz. 2021. Embracing Four Tensions in Human-Computer Interaction Research with Marginalized People. *ACM Trans. Comput.-Hum. Interact.* 28, 2, Article 14 (April 2021), 47 pages. https://doi.org/10.1145/3443686

[121] Maxim Lisnic, Alexander Lex, and Marina Kogan. 2024. "Yeah, this graph doesn't show that": Analysis of Online Engagement with Misleading Data Visualizations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–14. https://doi.org/10.1145/3613904.3642448

[122] Maxim Lisnic, Cole Polychronis, Alexander Lex, and Marina Kogan. 2023. Misleading Beyond Visual Tricks: How People Actually Lie with Charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–21. https://doi.org/10.1145/3544548.3580910

[123] Sophia B Liu, Leysia Palen, Jeannette Sutton, Amanda L Hughes, and Sarah Vieweg. 2008. In Search of the Bigger Picture: The Emergent Role of On-Line Photo Sharing in Times of Disaster. (2008).

[124] Colleen Long, Mike Balsamo, Michael R. Sisak, and Mary Clare Jalonick. 2024. Multiple Investigations and Security Failures in Trump Rally Shooting. https://apnews.com/article/trump-shooting-rally-a1bcbaa4a604e09be8cc22893751895b Accessed: October 26, 2024.

[125] German Lopez. 2024. The Myth of Migrant Crime. *The New York Times* (July 2024). https://www.nytimes.com/2024/07/18/briefing/the-myth-of-migrant-crime.html

[126] Nina Lutz, Jordyn W. Padzensky, and Joseph S. Schafer. 2024. Working with Color: How Color Quantization Can Aid Researchers of Problematic Information. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing* (San Jose, Costa Rica) *(CSCW Companion '24)*. Association for Computing Machinery, New York, NY, USA, 275–282. https://doi.org/10.1145/3678884.3681863

[127] Jim Maddock, Kate Starbird, Haneen J. Al-Hassani, Daniel E. Sandoval, Mania Orand, and Robert M. Mason. 2015. Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 228–241. https://doi.org/10.1145/2675133.2675280

[128] Lena Mamykina, Drashko Nakikj, and Noemie Elhadad. 2015. Collective Sensemaking in Online Health Forums. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3217–3226. https://doi.org/10.1145/2702123.2702566

[129] Douglas S Massey. 2020. The Real Crisis at the Mexico-US Border: A Humanitarian and Not an Immigration Emergency. In *Sociological Forum*, Vol. 35. Wiley Online Library, 787–805.

[130] Hana Matatov, Mor Naaman, and Ofra Amir. 2022. Stop the [Image] Steal: The Role and Dynamics of Visual Content in the 2020 U.S. Election Misinformation Campaign. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 541 (Nov. 2022), 24 pages. https://doi.org/10.1145/3555599

[131] Julie Carr Smyth Mike Catalini and Bruce Shipkowski. 2024. Trump falsely accuses immigrants in Ohio of abducting and eating pets. AP News. https://apnews.com/article/haitian-immigrants-vance-trump-ohio-6e4a47c52b23ae2c802d216369512ca5 Accessed: 2024-10-28.

[132] Daniel Moreira, Aparna Bharati, Joel Brogan, Allan Pinto, Michael Parowski, Kevin W Bowyer, Patrick J Flynn, Anderson Rocha, and Walter J Scheirer. 2018. Image provenance analysis at scale. *IEEE Transactions on Image Processing* 27, 12 (2018), 6109–6123.

[133] Brittany N. Morey. 2018. Mechanisms by Which Anti-Immigrant Stigma Exacerbates Racial/Ethnic Health Disparities. *American Journal of Public Health* 108, 4 (2018), 460–463. https://doi.org/10.2105/AJPH.2017.304266 arXiv:https://doi.org/10.2105/AJPH.2017.304266 PMID: 29470116.

[134] Michael Morisy. 2025. Federal Data is Disappearing. https://www.muckrock.com/news/archives/2025/feb/10/federal-data-is-disappearing-on-thursday-meet-the-teams-working-to-rescue-it-and-learn-how-you-can-help/

[135] Janice M Morse. 2007. Sampling in grounded theory. *The SAGE handbook of grounded theory* (2007), 229–244.

[136] Scott Neuman. 2024. Trump's close call: A detailed time line. *NPR* (July 2024). https://www.npr.org/2024/07/19/nx-s1-5041734/trump-shooting-assassination-crooks-bulter-secret-service

[137] Bryce Clayton Newell, Ricardo Gomez, and Verónica E Guajardo. 2017. Sensors, cameras, and the new 'normal'in clandestine migration: How undocumented migrants experience surveillance at the US-Mexico border. *Surveillance & Society* 15, 1 (2017), 21–41.

[138] Esme Nicholson. 2024. Rising anti-immigrant sentiment in Germany causes concerns in the country's business community. *NPR* (Aug. 2024). https://www.npr.org/2024/08/27/nx-s1-5050746/rising-anti-immigrant-sentiment-in-germany-causes-concerns-in-the-countrys-business-community

[139] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[140] David Noriega and Julia Ainsley. 2024. U.S. and China in high-level talks to deport more Chinese nationals, Mayorkas says. https://www.nbcnews.com/investigations/us-china-high-level-talks-deport-chinese-nationals-mayorkas-

says-rcna146590
[141] Milan Obaidi, Jonas Kunst, Simon Ozer, and Sasha Y. Kimel. 2022. The "Great Replacement" conspiracy: How the perceived ousting of Whites can evoke violent extremism and Islamophobia. *Group Processes & Intergroup Relations* 25, 7 (Oct. 2022), 1675–1695. https://doi.org/10.1177/13684302211028293 Publisher: SAGE Publications Ltd.
[142] Barack Obama. 2014. Presidential Memorandum – Response to the Influx of Unaccompanied Alien Children Across the Southwest Border. https://obamawhitehouse.archives.gov/the-press-office/2014/06/02/presidential-memorandum-response-influx-unaccompanied-alien-children-acr Accessed: 2024-10-27.
[143] Office of Ron Johnson. 2024. Border Security. https://www.ronjohnson.senate.gov/border-security Accessed: 2024-10-27.
[144] Office of the Director of National Intelligence. 2023. The Future of Migration - Global Trends 2040. https://www.dni.gov/index.php/gt2040-home/gt2040-deeper-looks/future-of-migration Accessed: 2024-10-27.
[145] Juan Sebastian Olier and Camilla Spadavecchia. 2022. Stereotypes, disproportions, and power asymmetries in the visual portrayal of migrants in ten countries: an interdisciplinary AI-based approach. *Humanities and Social Sciences Communications* 9, 1 (Nov. 2022), 1–16. https://doi.org/10.1057/s41599-022-01430-y Number: 1 Publisher: Palgrave.
[146] Committee on Pediatric Emergency Medicine. 2004. Overcrowding Crisis in Our Nation's Emergency Departments: Is Our Safety Net Unraveling? *Pediatrics* 114, 3 (09 2004), 878–888. https://doi.org/10.1542/peds.2004-1287 arXiv:https://publications.aap.org/pediatrics/article-pdf/114/3/878/1027716/zpe00904000878.pdf
[147] Graham C. Ousey and Charis E. Kubrin. 2018. Immigration and Crime: Assessing a Contentious Issue. *Annual Review of Criminology* 1, Volume 1, 2018 (Jan. 2018), 63–84. https://doi.org/10.1146/annurev-criminol-032317-092026 Publisher: Annual Reviews.
[148] Leysia Palen and Kenneth M Anderson. 2016. Crisis informatics—New data for extraordinary times. *Science* 353, 6296 (2016), 224–225.
[149] Leysia Palen and Sophia B. Liu. 2007. Citizen communications in crisis: anticipating a future of ICT-supported public participation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 727–736. https://doi.org/10.1145/1240624.1240736
[150] Leysia Palen, Sarah Vieweg, Jeannette Sutton, Sophia B Liu, and Amanda Hughes. 2007. Crisis informatics: Studying crisis in a networked world. In *Proceedings of the third international conference on E-Social Science*. 7–9.
[151] Orestis Papakyriakopoulos and Ethan Zuckerman. 2021. The Media During the Rise of Trump: Identity Politics, Immigration,"Mexican" Demonization and Hate-Crime. *Proceedings of the International AAAI Conference on Web and Social Media* 15, 1 (May 2021), 467–478. https://doi.org/10.1609/icwsm.v15i1.18076
[152] Britt Paris and Joan Donovan. 2019. Deepfakes and cheap fakes. (2019).
[153] Joon Sung Park, Michael S. Bernstein, Robin N. Brewer, Ece Kamar, and Meredith Ringel Morris. 2021. Understanding the Representation and Representativeness of Age in AI Data Sets. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 834–842. https://doi.org/10.1145/3461702.3462590
[154] Pew Research Center. 2024. About Pew Research Center. https://www.pewresearch.org/about. Accessed: 2024-10-28.
[155] Shruti Phadke and Tanushree Mitra. 2020. Many Faced Hate: A Cross Platform Study of Content Framing and Information Sharing by Online Hate Groups. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376456
[156] Whitney Phillips. 2018. The oxygen of amplification. *Data & Society* 22 (2018), 1–128.
[157] Etienne Piguet, Antoine Pécoud, and Paul de Guchteneire. 2011. Migration and Climate Change: An Overview. *Refugee Survey Quarterly* 30, 3 (2011), 1–23. https://doi.org/10.1093/rsq/hdr006
[158] Volkmar Pipek, Leysia Palen, and Jonas Landgren. 2012. Workshop summary: collaboration & crisis informatics (CCI'2012). In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion (CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 13–14. https://doi.org/10.1145/2141512.2141520
[159] Stephen Prochaska, Kayla Duskin, Zarine Kharazian, Carly Minow, Stephanie Blucker, Sylvie Venuto, Jevin D. West, and Kate Starbird. 2023. Mobilizing Manufactured Reality: How Participatory Disinformation Shaped Deep Stories to Catalyze Action during the 2020 U.S. Presidential Election. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 140 (April 2023), 39 pages. https://doi.org/10.1145/3579616
[160] Stephen Prochaska, Julie Vera, Douglas Lew Tan, and Kate Starbird. 2024. Misleading Ourselves: How Disinformation Manipulates Sensemaking. https://doi.org/10.48550/arXiv.2410.14858 arXiv:2410.14858.
[161] Lubna Razaq and Sucheta Ghoshal. 2024. What to the Muslim is Internet search: Digital Borders as Barriers to Information. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–17. https://doi.org/10.1145/3613904.3642770
[162] Kristina Fullerton Rico. 2025. Deportation fears create ripple effects for immigrants and their communities. *The Conversation* (2025). https://theconversation.com/deportation-fears-create-ripple-effects-for-immigrants-and-their-

communities-248817
[163] Graham H Ryan. 2024. Generative AI Will Break the Internet: Beyond Section 230. *Notes* (2024).
[164] Joni Salminen, Soon-gyo Jung, Shammur Chowdhury, and Bernard J. Jansen. 2020. Analyzing Demographic Bias in Artificially Generated Facial Pictures. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382791
[165] Joseph S. Schafer and Kate Starbird. 2023. Post-Spotlight Posts: The Impact of Sudden Social Media Attention on Account Behavior. In *Computer Supported Cooperative Work and Social Computing*. ACM, Minneapolis MN USA, 273–277. https://doi.org/10.1145/3584931.3606984
[166] Joseph S. Schafer, Kate Starbird, and Daniela K. Rosner. 2023. Participatory Design and Power in Misinformation, Disinformation, and Online Hate Research. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1724–1739. https://doi.org/10.1145/3563657.3596119
[167] Morgan Klaus Scheuerman, Alex Hanna, and Emily Denton. 2021. Do Datasets Have Politics? Disciplinary Values in Computer Vision Dataset Development. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–37.
[168] Morgan Klaus Scheuerman, Jialun Aaron Jiang, Casey Fiesler, and Jed R. Brubaker. 2021. A Framework of Severity for Harmful Content Online. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 368 (Oct. 2021), 33 pages. https://doi.org/10.1145/3479512
[169] Dietram Scheufele. 1999. Framing as a theory of media effects. *Journal of Communication* 49, 1 (1999), 103–122. https://doi.org/10.1111/j.1460-2466.1999.tb02784.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-2466.1999.tb02784.x.
[170] Norbert Schwarz and Madeline Jalbert. 2020. When (Fake) News Feels True: Intuitions of truth and the acceptance and correction of misinformation. In *The Psychology of Fake News*. Routledge. Num Pages: 17.
[171] Hyunjin Seo. 2014. Visual Propaganda in the Age of Social Media: An Empirical Analysis of Twitter Images During the 2012 Israeli–Hamas Conflict. *Visual Communication Quarterly* 21, 3 (2014), 150–161. https://doi.org/10.1080/15551393.2014.955501 arXiv:https://doi.org/10.1080/15551393.2014.955501
[172] Hyunjin Seo and Husain Ebrahim. 2016. Visual propaganda on Facebook: A comparative analysis of Syrian conflicts. *Media, War & Conflict* 9, 3 (2016), 227–251.
[173] De Rosal Igantius Moses Setiadi. 2021. PSNR vs SSIM: imperceptibility quality assessment for image steganography. *Multimedia Tools and Applications* 80, 6 (2021), 8423–8444.
[174] Michael D. Shear and Eileen Sullivan. 2024. Biden Confronts a New Border Crisis as Pressure Mounts from All Sides. *The New York Times* (2024). https://www.nytimes.com/2024/01/30/us/politics/biden-border-crisis-immigration.html Accessed: 2024-10-27.
[175] Irina Shklovski, Leysia Palen, and Jeannette Sutton. 2008. Finding community through information and communication technology in disaster response. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. ACM, San Diego CA USA, 127–136. https://doi.org/10.1145/1460563.1460584
[176] Robert Soden and Leysia Palen. 2018. Informating Crisis: Expanding Critical Perspectives in Crisis Informatics. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018), 162:1–162:22. https://doi.org/10.1145/3274431
[177] Kate Starbird. 2023. Facts, frames, and (mis)interpretations: Understanding rumors as collective sensemaking. *Center for an Informed Public* (December 6 2023). https://www.cip.uw.edu/2023/12/06/rumors-collective-sensemaking-kate-starbird/
[178] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 127 (Nov. 2019), 26 pages. https://doi.org/10.1145/3359229
[179] Kate Starbird and Leysia Palen. 2012. (How) will the revolution be retweeted? information diffusion and the 2011 Egyptian uprising. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) *(CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 7–16. https://doi.org/10.1145/2145204.2145212
[180] Kate Starbird and Leysia Palen. 2012. (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, Seattle Washington USA, 7–16. https://doi.org/10.1145/2145204.2145212
[181] Kate Starbird, Stephen Prochaska, and Ben Yamron. Conditional Acceptance with Minor Revisions, 2025. What is going on? An evidence-frame framework for analyzing online rumors about election integrity. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW (Conditional Acceptance with Minor Revisions, 2025).
[182] Kate Starbird, Emma Spiro, Isabelle Edwards, Kaitlyn Zhou, Jim Maddock, and Sindhuja Narasimhan. 2016. Could This Be True? I Think So! Expressed Uncertainty in Online Rumoring. In *Proceedings of the 2016 CHI Conference on*

Human Factors in Computing Systems (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 360–371. https://doi.org/10.1145/2858036.2858551

[183] Leo Graiden Stewart, Ahmer Arif, A. Conrad Nied, Emma S. Spiro, and Kate Starbird. 2017. Drawing the Lines of Contention: Networked Frame Contests Within #BlackLivesMatter Discourse. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 1–23. https://doi.org/10.1145/3134920

[184] Stefan Stieglitz, Deborah Bunker, Milad Mirbabaie, and Christian Ehnis. 2018. Sense-making in social media during extreme events. *Journal of Contingencies and Crisis Management* 26, 1 (2018), 4–15.

[185] Claudia Torrens and Elliot Spagat. 2024. As US debates immigration policy, migrant crisis worsens at Mexico's southern border. *Associated Press News* (2024). https://apnews.com/article/immigration-biden-border-title-42-mexico-asylum-be4e0b15b27adb9bede87b9bbefb798d Accessed: 2024-10-27.

[186] Alex Travelli. 2024. A Small Army Combating a Flood of Deepfakes in India's Election. *The New York Times* (June 2024). https://www.nytimes.com/2024/06/01/world/asia/india-election-deepfakes.html

[187] Jazmine Ulloa. 2024. Voting Rights Leaders Step Up Election Initiatives After Texas Raids. https://www.nytimes.com/2024/08/28/us/politics/voters-texas-raids.html

[188] Jazmine Ulloa and Nicholas Nehamas. 2024. Why Republicans Keep Calling Kamala Harris the 'Border Czar'. *The New York Times* (July 2024). https://www.nytimes.com/2024/07/17/us/politics/kamala-harris-border-czar.html

[189] Luisa Verdoliva. 2020. Media Forensics and DeepFakes: An Overview. *IEEE Journal of Selected Topics in Signal Processing* 14, 5 (2020), 910–932. https://doi.org/10.1109/JSTSP.2020.3002101

[190] Sarah Vieweg, Leysia Palen, Sophia B Liu, Amanda L Hughes, and Jeannette N Sutton. 2008. *Collective intelligence in disaster: Examination of the phenomenon in the aftermath of the 2007 Virginia Tech shooting*. University of Colorado Boulder, CO.

[191] Colin Ware. 2003. Design as applied perception. *HCI Models, Theories and Frameworks: Toward a Multidisciplinary Science. Morgan Kaufman Publishers: San Francisco (USA)* (2003).

[192] Karl Werder, Balasubramaniam Ramesh, and Rongen (Sophia) Zhang. 2022. Establishing Data Provenance for Responsible Artificial Intelligence Systems. *ACM Trans. Manage. Inf. Syst.* 13, 2, Article 22 (March 2022), 23 pages. https://doi.org/10.1145/3503488

[193] Tom Wilson, Kaitlyn Zhou, and Kate Starbird. 2018. Assembling Strategic Narratives: Information Operations as Collaborative Work within an Online Community. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 183 (Nov. 2018), 26 pages. https://doi.org/10.1145/3274452

[194] Robert Wolfe and Aylin Caliskan. 2022. American == White in Multimodal Language-and-Image AI. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Oxford, United Kingdom) *(AIES '22)*. Association for Computing Machinery, New York, NY, USA, 800–812. https://doi.org/10.1145/3514094.3534136

[195] Zhe Xu and Mengrong Zhang. 2023. How news media visually dehumanize victims of humanitarian crises through framing disparities: A quantitative comparative analysis. *International Communication Gazette* (Nov. 2023), 17480485231216583. https://doi.org/10.1177/17480485231216583 Publisher: SAGE Publications Ltd.

[196] Himanshu Zade, Megan Woodruff, Erika Johnson, Mariah Stanley, Zhennan Zhou, Minh Tu Huynh, Alissa Elizabeth Acheson, Gary Hsieh, and Kate Starbird. 2023. Tweet Trajectory and AMPS-based Contextual Cues can Help Users Identify Misinformation. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 103 (April 2023), 27 pages. https://doi.org/10.1145/3579536

[197] Brandy Zadrozny and Lora Kolodny. 2025. How Elon Musk boosted false USAID conspiracy theories to shut down global aid. *CNBC* (2025). https://www.nbcnews.com/politics/doge/elon-musk-boosted-false-usaid-conspiracy-theories-global-aid-rcna190646 Accessed: 2025-03-01.

[198] Xu Zhang and Lea Hellmueller. 2017. Visual framing of the European refugee crisis in Der Spiegel and CNN International: Global journalism in news photographs. *International Communication Gazette* 79, 5 (Aug. 2017), 483–510. https://doi.org/10.1177/1748048516688134 Publisher: SAGE Publications Ltd.

[199] Kaitlyn Zhou, Tom Wilson, Kate Starbird, and Emma S. Spiro. 2023. Spotlight Tweets: A Lens for Exploring Attention Dynamics within Online Sensemaking During Crisis Events. *Trans. Soc. Comput.* 6, 1-2 (June 2023), 2:1–2:33. https://doi.org/10.1145/3577213

10 Appendices

A Keywords

Below we provide a list of key phrases and hashtags that were run during this time of collection. Note that many of these are run as hashtags AND as phrases to collect data from Junkipedia. We use () to signify when we look up a term with and without that parenthesized word – i.e. "(biden) border crisis" represents looking up "biden border crisis" and just "border crisis".

- (biden) border crisis
- (biden) border invasion
- nyc migrants
- sanctuary city
- open borders
- close the border
- migrant crisis
- migrantes latinos
- border security
- migracion
- crisis fronteriza
- crisis migratoria
- texas border
- mass deportation
- chinese migrants
- cartel migrants
- trump border
- migrant invasion
- asylum seekers
- illegal immigration
- migrant bus
- border patrol migrants
- migrant shelters
- migrant EBT
- free migrant flights
- mass immigration
- illegal aliens
- colorado sanctuary city
- chicago sanctuary city
- migrant hotels
- migrant welfare

B Breakdown of data and data visualization sources

Below is a more detailed breakdown of our data visualization origins in this paper. We provide aggregate and disaggregate results. These results have undergone anonymization.

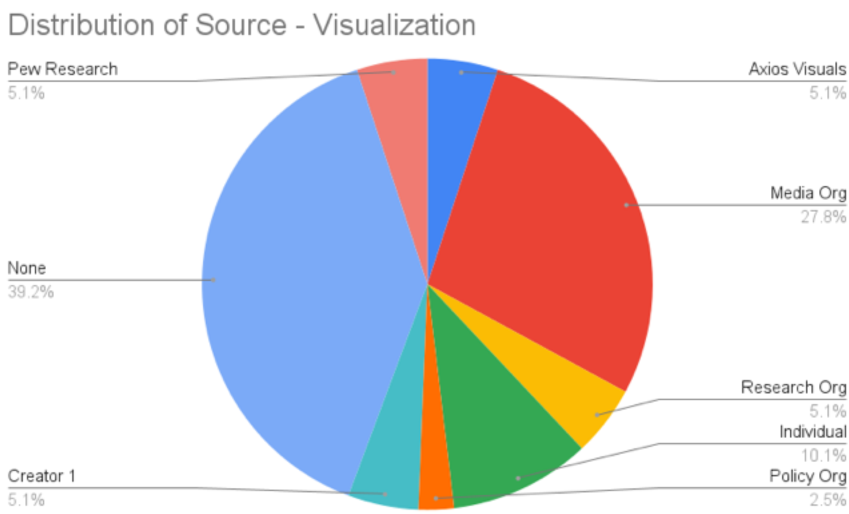

Fig. 33. Pie chart showing aggregated sources of data visualizations.

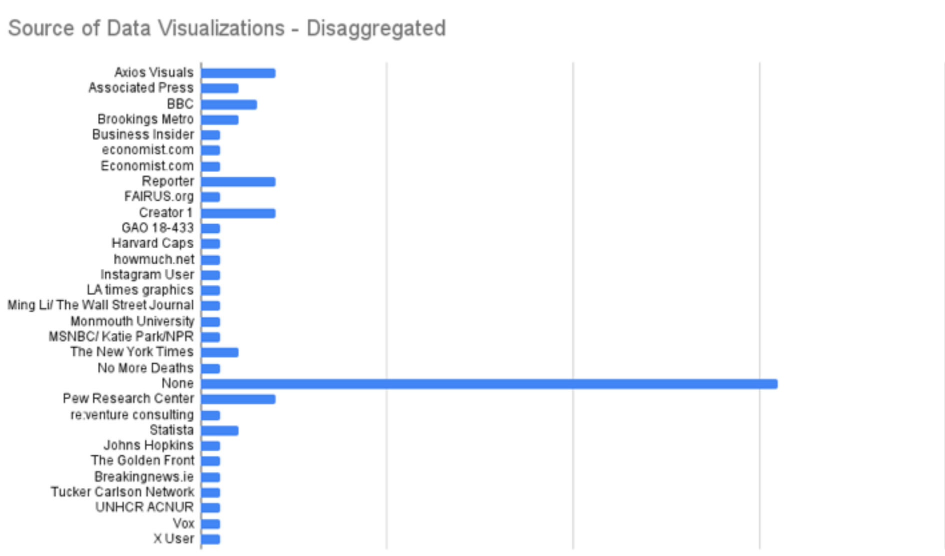

Fig. 34. Bar chart showing disaggregated sources of data visualizations. Anonymization has been performed on some of these sources.

Below is a more detailed breakdown of the origins of the data sources that power these visualizations in this paper. We provide aggregate and disaggregate results. These results have undergone anonymization.

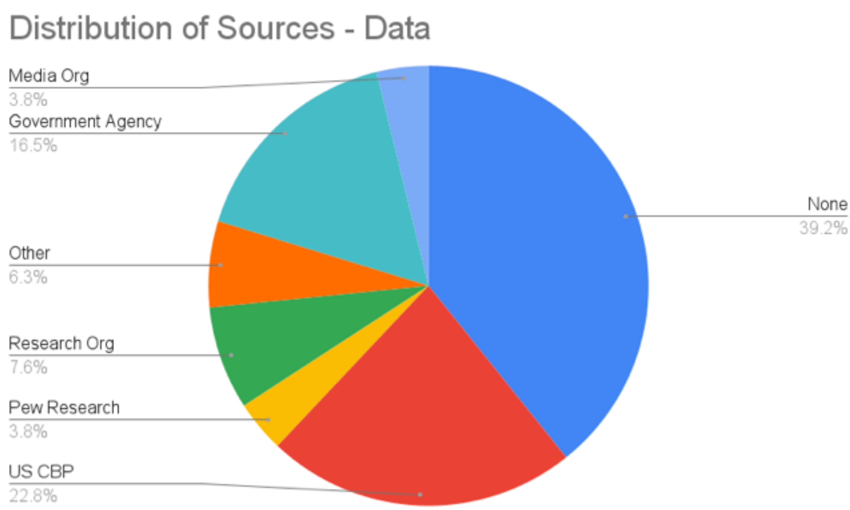

Fig. 35. Pie chart showing aggregated sources of the data behind visualizations.

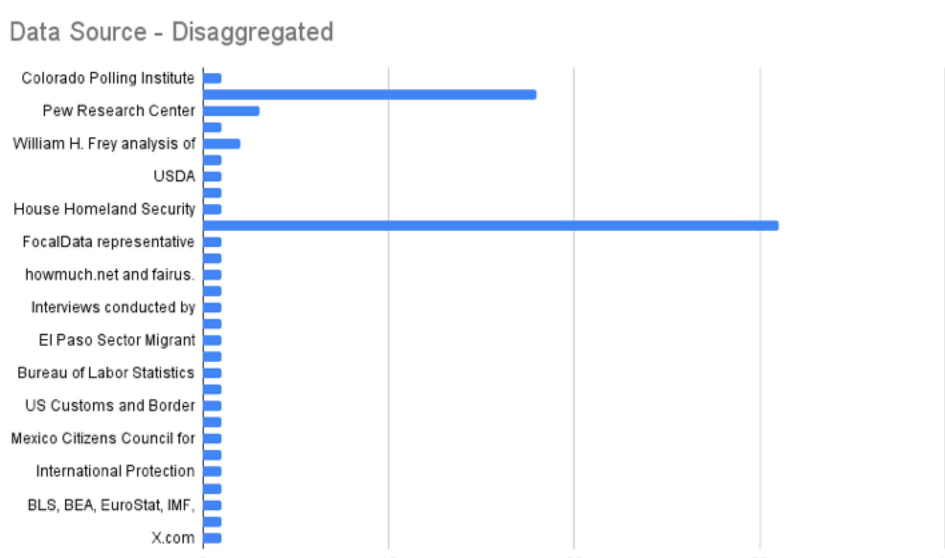

Fig. 36. Bar chart showing disaggregated sources of the data behind the visualizations. Anonymization has been performed on some of these sources.

C Codes breakdowns

Below we list our Cohen's Kappas by code:

Frame alignment - 0.716

Data units (X-Axis) - 0.783
Data units (Y-Axis) - 0.791
Visual Characteristics:
Visualization type - 0.715
Map type - 0.781
Chart type - 0.681
Visualization source - 0.732
Data source - 0.759
Annotation - 0.734
Watermark/Logo - 0.772

We do not report or measure"Other units" as it ended up not be applicable to enough visualizations in our sample.

Received 29 October 2024