

## 4 RESEARCHING

It was 2018, and the group Mumalá in Argentina had been documenting femicide for three years. Mumalá is the short name of the organization Mujeres de la Matria Latinoamericana (Women of the Latin American Motherland). The group was formed in 2001, during a period of economic crisis and great social need in Argentina. Mumalá characterizes itself as federal, feminist, popular, and dissident. The organization works across urban, rural, and remote parts of the country on gender rights, especially those issues that intersect with neoliberal economic violence. Mumalá's lead organizers started a femicide observatory in 2015, following the #NiUnaMenos uprising. At the time, the Argentine government had no central database of femicides (and it still does not have one that activists consider comprehensive).<sup>1</sup> Mumalá saw a need for independent monitoring that integrated a feminist analysis of the issue with specific political demands, and so it initiated its registry. Over three years the organization had amassed an extensive, centralized database of hundreds of cases, carefully researched and recorded by their members. But in early 2018, those same members walked away over political disagreements with the rest of the group—leaving the organization and taking the database with them.

Losing their data was a wake-up call for Mumalá. The organization used that painful moment to regroup and reformulate the way the group researched cases and stored data, as well as to expand its focus to monitor other forms of violence in addition to femicide. In late 2018, the organization relaunched what is now called the Mumalá Observatory: Women, *Disidencias*, Rights.<sup>2</sup> Members research and record cases of femicide, trans/travesticidios, femicide attempts, missing women, LGBTIQ+ hate crimes, and suicides of femicide perpetrators. As much as possible, they also try to record cases of indirect femicide, such as deaths from unsafe abortions.

Their research structure is now *federated*, meaning they have Mumalá affiliates based across Argentina's twenty-four provinces who monitor those specific locales for cases. This is better aligned with Mumalá's organizational structure, which is based on chapters that self-organize and coordinate local protests for gender rights around the country (see figure 4.1a). Thirty-two people in total produce data for the observatory. Like the majority of data activists we interviewed, their primary source of information is media reports. Mumalá members—whom they refer to as *compañeras* and *compañeres*—scan news media reports, conduct web searches, monitor Google alerts, and scan social media in order to discover new cases in the province they are responsible for.<sup>3</sup> They may also seek out official sources of information and work with their network of feminist journalists, who have more success than activists in speaking directly with prosecutors and judicial officials. Following the data loss in 2018, there is now not a single database but rather many multiple copies of the database. Activists work with their own local copies and then synchronize cases periodically. Mumalá publishes monthly and yearly reports to their social media accounts, often accompanied by analysis and infographics like the one in figure 4.1b.

The subject of this chapter is the work of *researching*—how data activists and journalists seek, find, and verify cases of femicide and related information, especially in the absence of official data. Researching is the second stage of a restorative/transformational data science project about femicide (see figure 2.4), and the researching process is at the heart of counterdata production work. Groups' analysis of power (from the *resolving* stage) is sharpened by encountering missing data from media and institutions during the researching stage. And *recording* cases is always a back-and-forth process with researching. As activists follow cases, sometimes over the course of months, years, or even decades, they continually seek and find new information and proceed to log that into more complete cases in their databases. As we will see, activists develop deep expertise in the information available in their contexts and engage in a vast array of creative informatic strategies to research missing data.

Compared to the groups we interviewed, Mumalá's research staff is quite large. Indeed, many femicide data projects are conceived and run primarily by individuals. Such is the case with Dawn Wilcox, who runs Women Count USA from her home in Texas. She has a full-time job as a school nurse, but she spends many hours on nights and weekends scanning news media articles and following up on tips sent by strangers. Her goal is to document every case of femicide in the United States going back to 1950. A survivor of domestic violence herself, Wilcox became curious about femicide statistics in the United States in 2016. She started what she thought was a simple online information search and was shocked when she realized how difficult it was to



## 221 FEMICIDIOS



Observatorio  
**mumala**  
MUJERES DE LA MATRIZ LATINOAMERICANA

2 de 10

FIGURE 4.1

(a) Mumalá protesting on the seventh anniversary of #NiUnaMenos in June 2022. Courtesy of Mumalá. (b) Mumalá’s observatory registered 221 femicides in 2021, including 6 transfemicides/travesticides. Courtesy of Mumalá.

find systematized information on women's murders. She found that the FBI's Uniform Crime Report was woefully incomplete, not only for women but for all homicides, as it relies on data voluntarily reported from law enforcement agencies. Other citizen-run archives, like the National Gun Violence Memorial, only register gun deaths. And domestic violence organizations tended to limit their scope to intimate partner relationships, which left out women killed by other family members, by neighbors, by strangers, in the context of sex work, and more. Wilcox felt that the fragmentation of the information contributed to the invisibility of the problem: "I felt like if I could bring all of this data into one place that, first of all, it could tell a story about what was really happening to women and where it was happening, how it was happening, who was doing it, who was killing women, what sort of relationships they had. And I felt like it would memorialize these women, which was very important. It would show that they were more than just statistics on a page."

Wilcox sources cases of femicide daily from digital news articles that she finds by using a search engine and typing a set list of queries such as "woman's body found" or "husband kills wife". Since Wilcox started the project, her work has become more widely known, and people will often email her news articles about femicide. Like many activists that we spoke with, she cross-references or "triangulates" multiple news media articles against each other and with other sources of data in order to verify information and arrive at the details she needs to log a case in her database. Reading a news article will reveal a victim's name, which she will then use to search for more details. She is adamant that every woman in her database needs a photo, and this is the piece of information that takes the longest amount of time to find. When the photos that she finds are low quality, she will retouch them. The work is immense: "I think even if I did this work full time, I would still need help. It's just . . . the sheer number of cases is just staggering." Wilcox is always working with a backlog of articles to review, cold cases to search up, and tips to follow up on.

All activists face challenges in the researching stage, but groups that monitor MMIWG2 and racialized feminicide face even more hurdles. News outlets cannot be their sole source of discovering new cases because they know that the media systematically neglect to report on this violence. Such is the case with the Sovereign Bodies Institute, introduced in the prior chapter, which monitors MMIWG2 and MMIP across the Americas and beyond. Official data are woefully inadequate, and Sovereign Bodies Institute has quantified exactly how inadequate they are. In a scan of official records in California, Sovereign Bodies Institute found that 91 percent of missing Native girls in California are also missing from at least one official database. Of those actually recorded in state databases, 56 percent of missing Native women and girls in California

were classified incorrectly as a different race.<sup>4</sup> Government and law enforcement claim that part of the problem is jurisdiction confusion—what they refer to as *jurisdictional mazes*—meaning that it is unclear to officials themselves which county, state, federal, or tribal agency is responsible for investigating a crime. Sovereign Bodies Institute and other groups refute that claim by creating clear flow charts such as that in figure 4.3 and educating families about jurisdiction.<sup>5</sup> For Sovereign Bodies Institute, trying to detect cases from media articles is challenging because the media, first, don't report on MMIWG2 cases and, second, even when they do consider those cases newsworthy, according to Annita Lucchesi, founder and director of Research + Outreach, the “press don't really do a great job of acknowledging victim Indigeneity.”

Instead of relying on official data or news media, then, Lucchesi describes how Sovereign Bodies Institute researchers “get creative with the data.” They use Indigenous networks, both digital and analog, to discover and gather information about new cases. They review social media posts and do direct outreach. They have done public records requests through the Freedom of Information Act (FOIA), partnered with tribal enrollment offices, and used historical archives. As a family-centered, survivor-led, and survivor-centered organization, Sovereign Bodies Institute pairs its informational work with direct services to families and tries to respond to all family requests for help on specific cases. At the time of our interview in 2020, Lucchesi was working directly with a medical examiner for the state of Montana to carefully review state autopsies for a handful of cases that families wanted reopened, challenging the state ruling of “suicide” as a misclassification. Sovereign Bodies Institute's sources of information end up needing to be “multipronged” and “diverse” to counter the additional research hurdles faced by groups monitoring violence against Indigenous women.

## RESEARCHING CASES

These three groups—Mumalá, Women Count USA, and Sovereign Bodies Institute—use diverse research strategies to find information about femicide in their contexts. Researching is arguably the most time-consuming part of any counterdata production effort. This is where an individual or group spends the bulk of their time and effort: seeking, finding, and verifying cases of femicide. The researching stage includes discovery of new cases as well as ongoing research to follow, add to, and verify information for other existing cases (see figure 4.2). During this stage of a restorative/transformational data science project, activists continually assess existing sources of information, including official databases, other counterdata and citizen data projects, news media, social media, and more.

# RESEARCHING

## Finding + verifying information

Activists seek relevant information to add to their database. This can include sourcing existing datasets, mining media and other sources of information, and triangulating across sources to verify details. Such research either discovers new cases or adds information to existing cases in the database.

### RESOLVING

Developing a theory of change

### RESEARCHING

Finding + verifying information

### RECORDING

Information extraction + classification

### REFUSING + USING

Where data go, who uses them

In their research process, all groups encounter many examples of missing data—data that are either wholly absent or else incomplete, incorrect, biased, or inaccessible. Missing data can function as both a motivating factor for beginning a project and as an ongoing challenge to navigate as activists research cases. For Mumalá, it was the failure of the government to create a federal registry of femicide in Argentina, even after the #NiUnaMenos movement had made that a central demand and the government had committed to fulfilling it. For Wilcox, it was the incomplete and underreported federal homicide database combined with civil society data projects focused on subsets of femicide, but not femicide itself. And in Sovereign Bodies Institute’s case, they encounter incomplete and misclassified official databases, law enforcement who use “jurisdictional mazes” as an excuse to not investigate, and heightened media bias, making cases of MMIWG2/MMIP especially challenging to document.

All activists and journalists doing this work are deeply concerned with getting the facts and details of a case straight. They learn to be suspicious of early media reports about a femicide because these can often be filled with errors. They typically need to follow a case for some time until details are clarified. All groups seek to use multiple sources of information about a given case of femicide to verify details. Once they obtain the name of the victim or some identifying details, activists can search for more media articles about that case as well as seek out institutional data such as police reports, court records, or death registries, a process they called *triangulation*. For certain groups—such as the groups described in chapter 2, the Fórum Cearense de Mulheres in Brazil, and the Grupo Guatemalteco de Mujeres—advocates start with the official data and then triangulate that with media articles in order to understand case details, context, and whether the state had properly classified the death. Activists learn which outlets cover the issue, how it is reported, and which terms and language and framing are used. For example, Femicidio.net is a group that has been monitoring femicide in Spain and producing reports since 2010. Given over a decade of experience navigating different sources of information, they have developed extensive expertise in the Spanish digital media ecosystem. Nerea Novo, a researcher with the group, says it like this: “The methodology [for searching for cases] can be taught quickly but then practice



FIGURE 4.2

Researching is the second stage of a restorative/transformational data science project. Courtesy of the author. Graphic design by Melissa Q. Teng. Collaged images: Courtesy of Ana María Abruña Reyes / Todas (<https://todaspr.com>).

gives you certain tricks to get to know which sources are more reliable, which sources less, which sources dedicate more time to research and which do not.”

The work of researching cases—reading and scanning stories of violence, sometimes for hours every day—takes an emotional and mental toll on data activists. Betiana Cabrera Fasolis, from Mumalá, concedes, “Yes it’s true . . . the work we do is quite pessimistic. The compañeras never cease to be surprised at the amount of violence in new cases and old.” Researching femicide involves the continual emotional labor of reading about brutal violence and the secondary witnessing of the trauma and loss of others, a question we will return to at the end of this chapter. For survivors of gender-related violence, it is work that hits close to home.

#### POWER AND RESISTANCE IN THE INFORMATION ECOSYSTEM

In undertaking research, activists operate in an *information ecosystem*: a dynamic constellation of actors that includes infrastructure, tools, technology, producers, consumers, curators, and sharers of information about femicide. The metaphor of the ecosystem is designed to capture the dynamic nature of information; it moves and flows across scales and sites and actors as it is produced, curated, transformed, and used.<sup>6</sup> In the case of femicide information, the sites and actors that surfaced most frequently in our interviews included the state and the media as key producers of information about cases (and, paradoxically, also key drivers of missing and biased information). Families themselves also provide information to activist groups.

As we interviewed activists and journalists, they provided analysis and insider perspective on their information ecosystems. They have expert answers to questions like the following: Who produces information about femicide cases? How reliable is it? How can it be verified? How public is it? If it is not public, what are other ways to acquire it? When is information biased or unreliable? Why and how is information missing, biased or dubious? Indeed, we found that activists connect their analysis of power (discussed in the prior chapter on resolving) into concrete observations about how that power is made manifest in the information ecosystem in their country or locale, and they then use creative strategies to address those informatic gaps and biases.

Table 4.1 places activist analyses of missing data in counterpoint to the strategies they use to overcome that missing data through research.<sup>7</sup> Activist analyses of power are grouped into the domains that surfaced most frequently: the state (including both laws and the implementation of laws), media, and families. Each of these domains is important in the information ecosystem because each significantly affects the production of data and information about femicide. These groupings also map to the



Table 4.1

Activist analyses of missing data and research tactics they use to overcome it

Domain	Reasons for missing data mentioned by activists	Research tactics activists use to overcome missing data
<b>State—laws</b>	<ul style="list-style-type: none"> <li>• Absence of laws on femicide</li> <li>• Narrowness of laws on femicide</li> <li>• No public disclosure laws</li> <li>• Law doesn't recognize certain groups (trans people, Indigenous status of people)</li> <li>• Fragmentation of laws</li> </ul>	<ul style="list-style-type: none"> <li>• Include cases based on structural analysis of power</li> <li>• Count people that are excluded by present laws, including LGBTQ+, children, Indigenous people not recognized by state</li> <li>• Use nonstate information sources</li> </ul>
<b>State—implementation of laws</b>	<ul style="list-style-type: none"> <li>• Narrow interpretation of law</li> <li>• State ignores or avoids disclosure requirements</li> <li>• Lack of state resources to investigate cases and/or publish information</li> <li>• State employee turnover (police, prosecutor, judicial)</li> <li>• Lack of state expertise in gender, race, femicide</li> <li>• Public information not disaggregated (by gender, race, sexuality)</li> <li>• Public information shows signs of political manipulation</li> <li>• State is absent in rural/remote areas</li> <li>• State consistently misclassifies cases (e.g. suicide cases, cases related to Indigenous women, trans people)</li> <li>• Information collection and reporting is fragmented across state agencies</li> </ul>	<ul style="list-style-type: none"> <li>• Manually review state websites, state social media feeds, state-run WhatsApp groups</li> <li>• Search court records</li> <li>• Follow up on individual misclassification/suicide cases, especially where family and community are contesting state ruling</li> <li>• Triangulate state records with media reports to see what's conflicting</li> <li>• Visit morgues and medical examiners</li> <li>• Work in networks and coalitions to discover cases and share information</li> <li>• Solicit/receive crowdsourced reports of cases</li> <li>• Use federated media monitoring structure to cover large territories</li> <li>• Use media sources—TV, radio, print, digital—to discover cases</li> <li>• Use other citizen-led databases</li> </ul> <p>* Partner with tribes and cross-reference their records                      * Mine state-run historical archives                      * File public records requests, like FOIA in the United States                      * Call friends who work for the state and discuss discrepancies</p>

Table 4.1 (continued)

Domain	Reasons for missing data mentioned by activists	Research tactics activists use to overcome missing data
<b>Media</b>	<ul style="list-style-type: none"> <li>• Media report sensational details but do not provide info on race, tribe, sexuality</li> <li>• Media report on some deaths and not others (e.g. often don't report on trans, Black, Indigenous, migrant, rural/remote, poor people)</li> <li>• Media reporting about femicide is biased, toxic and sensational</li> <li>• Media draw from police reports and framings, replicating state biases (like misgendering trans people) and use victim-blaming frame</li> <li>• Media are absent in rural/remote areas</li> <li>• Media do not report on killings related to organized crime/paramilitary activity for fear of retribution and violence</li> <li>• Media do not follow cases through the justice system</li> <li>• Early media reports on killings are full of incorrect information</li> </ul>	<ul style="list-style-type: none"> <li>• Use stigmatizing language to search for cases in media, like “crime of passion,” “man dressed as woman,” but reject their framing</li> <li>• Seek cases from hyperlocal and regional news outlets and blogs</li> <li>• Use social media (esp. Facebook) and private groups (esp. WhatsApp) to find and validate cases outside mainstream media</li> <li>• Use networks and partnerships to confirm details/verify information</li> <li>• Collect humanizing information (e.g., photos, details about life)</li> <li>• Triangulate state records with media reports to see what's conflicting</li> <li>• Follow individual cases through judicial system</li> </ul> <p>* Use death certificates to corroborate race (with caution, activists know these are often wrong)</p> <p>* Follow prominent journalists reporting on femicide</p>
<b>Families</b>	<ul style="list-style-type: none"> <li>• Family/community doesn't report because too much work, already traumatized</li> <li>• Stigmatizing for families to report femicide (e.g., elite families try to keep cases out of the press)</li> <li>• Minoritized communities have good reasons not to trust the authorities so they don't report</li> <li>• Families may face violence for reporting</li> </ul>	<ul style="list-style-type: none"> <li>• Partner and share info with groups that provide services to families</li> <li>• Families/communities contact activists directly to request inclusion or to share more details</li> </ul> <p>* Get individual case files from families, friends, or state leaks</p> <p>* Contact families directly to verify information or offer support/services</p>

\* = Strategy mentioned by only 1–2 groups. Not a common pattern.

domains of oppression outlined in Patricia Hill Collins' matrix of domination, introduced in chapter 2. While that chapter provided a high-level analysis of different factors at play in producing missing data about femicide in each domain, our interviews helped to elucidate activists' own analysis of factors that led to missing data in their locale.

#### MISSING DATA: STATE LAWS

Many interviewees mentioned either narrowly framed legislation or the complete absence of legislation leading to the lack of state-produced information (see table 4.1, first row). As Myrna Dawson from the Canadian Femicide Observatory for Justice and Accountability stated, "In Canada, nothing is officially seen as a femicide because we don't have any legislation or any official recognition of femicide." When laws do exist, they may be narrowly formulated. A handful of laws only include intimate partner violence, and many exclude or make no provisions for transgender women, so these cases will not be included in official counts, lists, or statistics (see table 1.1). Even when laws and official data do exist, activists mentioned that the absence of public records laws can inhibit the availability of information.

Activists counter these legislative hurdles by deliberately defining and counting femicide in a way that matches their own structural definition of the violence, which typically exceeds the state's legal definitions. For example, Mumalá counts induced suicides—when a woman is driven to suicide by repeated domestic abuse—as femicides. This is a concept not currently outlined in Argentine law but it is recognized in El Salvador's femicide law, which Mumalá has drawn inspiration from. They also count deaths from unsafe abortions as femicides. This information became a central part of the massive national movement for the legal right to abortion in Argentina in the years leading up to 2020, because it showed that unsafe abortions were a leading (and preventable) cause of maternal mortality. Mumalá also seeks information on transgender killings, *travesticidios*, and also on the premature deaths of trans/*travesti* people, which Mumalá attributes to gendered forms of social and economic violence that reduce trans life expectancy to half that of the cisgender population. By necessity, they must seek nonstate sources of information to register such violence. But any such expansion is only done with careful consideration. For example, in the absence of adequate legal definitions for phenomena like transfemicide, Mumalá relies on discussion, debate and consensus-building about individual cases through their WhatsApp group, what Cabrera Fasolis calls "the soul of the observatory." Through such collective deliberation, the group determines whether a case corresponds to their definition of femicide or trans/*travesticidio*, using their analysis of power, and therefore whether it should be counted.

#### MISSING DATA: STATE IMPLEMENTATION OF LAWS

Likewise, in terms of the implementation of laws, groups mentioned a variety of factors leading to state oversight, bias, and mismanagement of cases that affected downstream information systems (see table 4.1, second row). For example, activists and journalists highlighted the lack of state resources to implement laws, investigate cases, and/or publish information. For the Alianza Feminista para el Mapeo de los Femi(ni)cidios en Ecuador, among many others, they saw an acute absence of the state in rural areas of the country, leading activists to speculate that cases in those areas are much higher than official counts. Some groups mentioned the frequent turnover in law enforcement officials, leading to the loss of institutional information about cases. As Julliana de Melo from Uma por Uma noted, “Local police chiefs in the Pernambuco police change constantly which hinders the investigation process a lot. We saw the same case passed on to several local police chiefs who did not have access to the amount of data that we had. [. . .] So, this shuffling, this inconstancy on the part of the police, was also something that we realized was a failure in the system.”

Other activists surfaced questions of bias and training, noting how the public sector employees are not trained to recognize or investigate gender-related violence, leading to disregarding such factors in cases or to misclassify cases. State misclassification, especially of race/ethnicity (for Indigenous people), gender (for trans and gender nonconforming people), and cause of death (for accidental deaths, induced suicides, and police violence) was one of the most frequently mentioned drawbacks of official data. Even when the state may produce information on fatal gender-related violence, activists stated that there is often institutional fragmentation over which agency collects which cases, leading to multiple “counting” arms of the state that count different phenomena with different criteria. This is an observation corroborated by ILDA’s work on their data standard in Latin America, discussed in chapter 1, which found that states lack both technical capacity and shared methodology to share femicide data across divisions internally. Finally, many activists noted that official information about femicide is often not published in a timely and disaggregated way, and it occasionally shows signs of tampering, as was mentioned in the case of domestic violence data in Puerto Rico in chapter 2.

Despite challenges in acquiring official information, activists get resourceful to mine what they can from diverse state sources. For example, many manually review state websites, attorney general pages, court records, state-based social media feeds, or state-sponsored chat groups on WhatsApp or Telegram. These never yield “data” in the sense of systematized information in rows and columns, but activists can extract important details about individual cases for their databases. Activists based in diverse

places—Guatemala, the United States, and Canada—have also resorted to visiting morgues to interview state employees and review specific cases together. For example, Sovereign Bodies Institute’s aforementioned work with a medical examiner in Montana led to a case of accidental death being reexamined as a murder. As Lucchesi recounted the story to us, she described the impact of the collaboration: “It was very graphic and traumatic at first but it was also really empowering. And I was able to explain things to the family that I wouldn’t have been able to explain otherwise.”

While it can be acquired creatively, information from the state may still be biased, misclassified, or unreliable, so activists have various strategies to triangulate it with nonstate information. These include incorporating cases from other counterdata or citizen-led projects. For example, in the United States, both Wilcox and the African American Policy Forum periodically copy relevant cases from the Fatal Encounters database, a grassroots effort that documents police violence. Lucchesi said she tried to do the same for Indigenous women, but Fatal Encounters had placed Native cases into the “other” racial category, illustrating how counterdata projects can themselves replicate the same biases present in official data sources.

Activists develop novel and collaborative human networks of information-sharing in order to fill in gaps in official data. This is evidenced by Mumalá’s federated monitoring structure where they are able to cover a vast geographic territory by having members responsible for discovering cases in each province. The Alianza Feminista para el Mapeo de los Femi(ni)cidios en Ecuador works in a similar way where their coalition includes regionally based groups that serve as a key source for information about cases in rural and remote areas. Many groups have developed relationships with feminist journalists and leverage those relationships to source information. And finally, once activists’ work becomes well-known, like Wilcox’s and María Salguero’s, they begin to receive many crowdsourced tips about cases through email, messaging apps, and social media.

#### MISSING DATA: MEDIA

By and large, the most prevalent source of nonstate information for femicide is the news media (see table 4.1, third row). Several times a week, members of Mumalá will type search queries such as “death of a woman Cordoba” into search engines to source recent news reports. They also plug those queries into Google Alerts so that they get notified when Google indexes new web pages that meet those criteria.<sup>8</sup> Because Mumalá distributes monitoring across individual provinces, the geographic modifier—“Cordoba”—is meant to limit the search to reports from the province of Cordoba. But search engines don’t always index the crime tabloids, small regional news sites, and

local blogs that activists find most helpful in their research. So compañeras and compañeres from Mumalá find they still have to go to the website of each news outlet in that province and read through different sections where feminicides may be reported. Most groups do this digitally, going site by site and section by section, but a small number of groups work in other formats. Members of the Grupo Guatemalteco de Mujeres read and clip physical newspapers, and Carmen Castelló watches the daily news on TV to discover new cases in Puerto Rico.

While all groups use news media as either a primary or secondary source for researching cases, all groups are also deeply critical of media reporting on femicide and gender-related killing. It was variously called sensational, irresponsible, shameful, victim-blaming, misgendering, dehumanizing by design, stigmatizing, toxic, transphobic, lesbophobic, racist, xenophobic, demeaning, trauma porn, misery porn, and poverty porn. Activists described how media often draw directly from police reports and quote law enforcement as an authoritative source, which leads to the transmission of bias—a kind of collusion between the state and the media. A story recounted to us by Toni Troop illustrates this point. At the time of our interview, Troop was the executive director of Jane Doe, an organization that monitors fatal domestic violence in the US state of Massachusetts. She described how she was deeply frustrated that news articles always quoted police and never reached out to domestic violence prevention organizations. She invited three local Boston reporters to a meeting and asked them directly why that was. “And that was the one of the most enlightening conversations of my career,” she emphasized, “because what they said was, ‘We’re going out to get the facts. And the facts change between our first story, versus the second story, until there’s an arraignment. Some of the police officers and the DA’s offices, they’re not even calling it domestic violence yet. It’s not up to us to do that.’” At that point, Troop had an aha moment in which she realized that the press felt unqualified to name the violence as *domestic violence* until a state official had uttered that term. While the press felt that they could not name the violence unless police named it, they did not consider the harm they might be doing by *not naming* the violence or by solely relying on police as the authoritative namers, framers, and definers of violence.

In many cases, activists have to adopt the stigmatizing language used by the media into their search queries in order to retrieve the articles they need (i.e., googling “crime of passion”). However, activist groups generally reject the framing of the news articles, which tend to report on killings in ways that sensationalize the violence, depict them as isolated incidents, and blame victims for their own deaths. This was why Cuántas Más ran so many workshops, as discussed in chapter 3, trying to improve journalists’ understanding of gender-related violence. This is to say that media articles are useful for extracting specific fields needed for activist databases—such as victim age, method

of death, or the relationship between perpetrator and victim—but they are toxic and harmful for the more important tasks of framing and analyzing the phenomenon. This is why groups’ development and use of their own analysis of power, described in the prior chapter, is so important.

In response to harmful media narratives, activists reject dominant media framings of cases, use social media, and work with organizations on the ground, and often directly with families, to verify important details about cases. Like Wilcox, many activists seek out humanizing information such as photos or details about people’s lives in order to give them some form of memory justice that the media has denied them. Where news media tend to overrepresent the point of view of the state, activists strive to believe, support, and follow the lead of victims’ families and communities. For example, La Casa del Encuentro’s observatory was founded in 2008 because the family of Adriana Marisel Zambrano, an Indigenous woman from a rural province in Argentina, could not get the attention of either the judicial system or media and reached out to the feminist organization for support.

Another built-in flaw in relying on media reports is that, in the words of Silvana Mariano from Néias—Observatório de Femicídios Londrina in Brazil, “the press are selective.” All groups recognize this as an inherent limitation of using news articles to systematically detect cases of femicide, but some groups experience it more acutely. Activists like Mumalá who monitor large geographic territories discussed how media don’t cover those areas or, in some rural and remote places, simply don’t exist. There are other regions where media do not report on violence related to narco-trafficking for fear of retaliation from organized crime networks. Groups that monitor gendered violence at the intersection of race, Indigeneity, and sexuality, such as MMIWG2, LGBTQ+ killings, and Black women killed in police violence, must look beyond news articles to discover cases because they are systematically underreported in the media, as discussed in chapter 2. These groups develop extremely creative strategies, such as the Sovereign Bodies Institute mining historical archives, doing FOIA requests, and developing partnerships with tribal enrollment offices. Gregory Bernstein, who worked at the African American Policy Forum on their database of Black women killed in police violence in the United States, stated that such work requires “finding different, inventive ways of learning these stories.” But at the same time, this work requires activists to acknowledge the incompleteness of their work; because the information is so challenging to find, Bernstein says, “we know that there are so many that we are missing.”

#### MISSING DATA: FAMILIES

While mentions of families and communities connected to missing data were less frequent, activists did note some situations that led to families not reporting cases

and thus to resulting gaps in information (see table 4.1, last row). First, while the pain of loss may lead some families to seek to make their cases public in order to secure justice, others handle their trauma by refusing to engage with either the state or media around the case. Numerous groups discussed that as *femicide* has grown in usage as a term, a stigma has been associated with it. This has led to wealthier families trying to suppress information about cases, particularly when the perpetrator is another member of the family. In contexts of racialized state violence like the United States or widespread corruption such as El Salvador, families and communities have good reasons to not trust the authorities and may not report for fear of incurring retribution or further trauma. For example, Kimberlé Crenshaw, legal scholar and organizer of the #SayHerName campaign, and her colleagues wrote about the case of Kayla Moore, a transgender woman in Berkeley, California, who was undergoing a mental health crisis.<sup>9</sup> Her roommate called the police who, instead of providing support, attempted to arrest her and ended up suffocating her to death in her own bedroom. As their paper, and the broader #SayHerName campaign, demonstrates, these stories of state violence represent a widespread and systematic pattern, leading minoritized communities to not report and to not engage with institutions that harm and terrorize them.

Working with families can be an important source of information for activists to confirm details about a case, but activists are careful in how they engage with families. Some, like Sovereign Bodies Institute, do provide services and support to families, but Lucchesi emphasized the importance of waiting for families to contact them: “We don’t reach out to families directly until they’re ready. That trauma is so severe that there are huge unintended consequences that can come from directly soliciting family.” The majority of data activists do not work directly with families, but rather offer various forms of *acompañamiento* (accompanying) to families and family-centered advocacy groups. This might mean providing space for their meetings, showing up for marches organized by family groups, providing data and information for family-led vigils, or other forms of support and solidarity. Through these relationships and partnerships, data activists and family groups often do end up sharing information with each other, sometimes even full case files. These can be important sources, especially when such information is not available from state or media sources.

#### INFORMATIC RESISTANCE IS CONTEXTUAL AND RELATIONAL

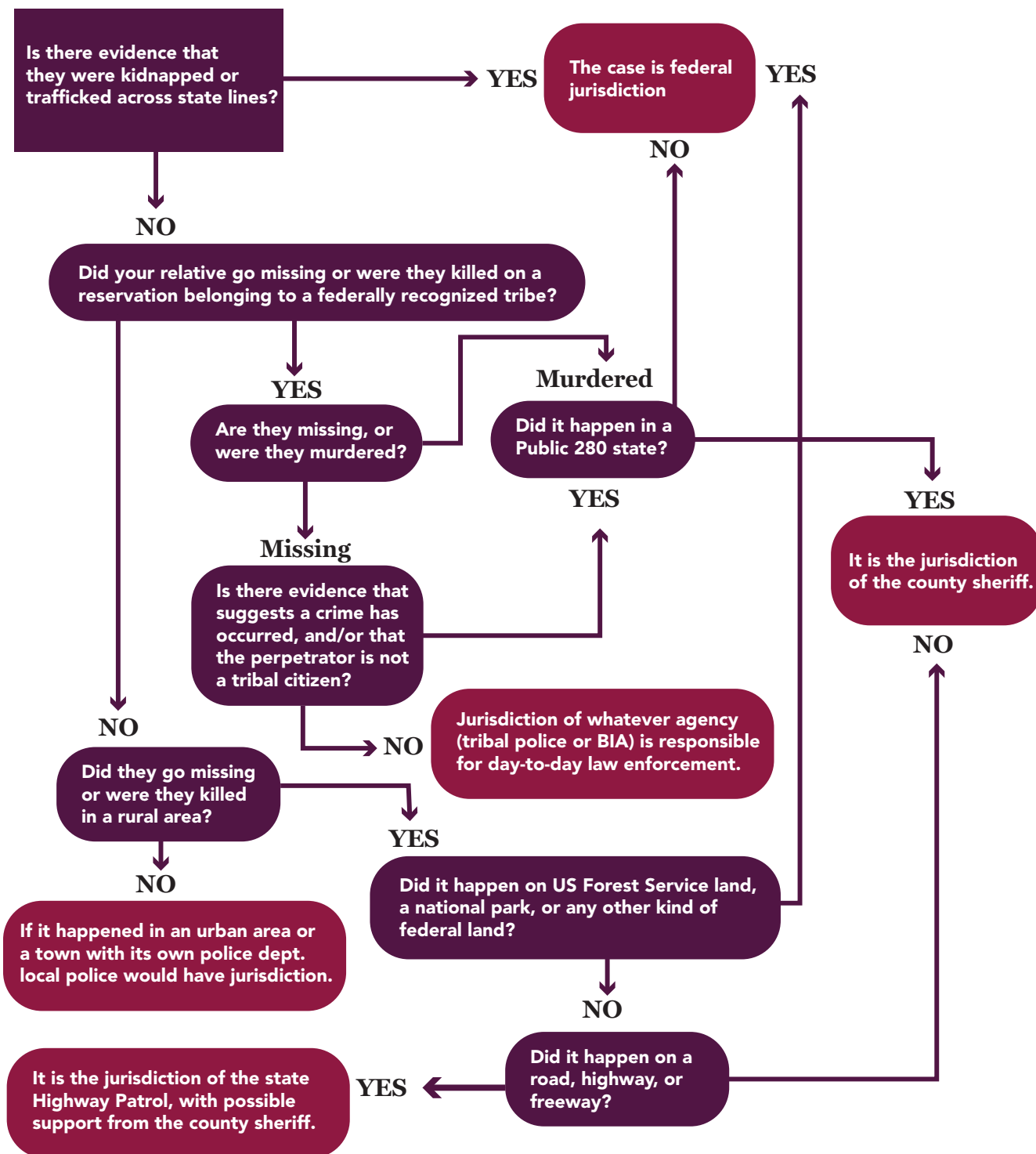
As a consequence of combining their power analysis with their researching practices, groups develop deep expertise in the flawed information ecosystems surrounding



femicide and gender-related killing. This is what Lauren Klein and I meant when we outlined the data feminism principle to *consider context*: “Data are not neutral or objective. They are the products of unequal social relations, and this context is essential for conducting accurate, ethical analysis.”<sup>10</sup> Femicide data can never be collected, compiled, nor used at face value, because the contextual conditions of their production so deeply affect their quality.

Activist expertise includes not only creativity and skill in locating individual data points to enter into their spreadsheets, but deep familiarity with the political, historical, legal, cultural, and geographic factors—the contextual factors—affecting the production, availability, and reliability of information about the issue. We can see this evidenced by Sovereign Bodies Institute’s MMIWG2 Jurisdiction Flowchart in figure 4.3, which they use to teach families and advocates about the legal geographies of a particular case (as well as to refute law enforcement’s claim that jurisdictional mazes are preventing them from launching an investigation). Sovereign Bodies Institute published the flowchart in their organizing toolkit in 2020, and it walks the viewer through a series of questions in order to determine legal jurisdiction of a particular case in the United States. Factors like whether the victim was taken over state lines, whether the crime happened on a highway, whether the crime happened on a reservation, and whether or not the perpetrator is a tribal citizen all shift which agency would be responsible for handling a case. Note that many of these details may not be known at the outset of a crime, so the responsible agency can shift while a case is under investigation, and information may be (often is) lost in the transfer. In-depth understanding of the legal landscape—as well as its shortcomings—has allowed Sovereign Bodies Institute to amass more complete and correct case data. In fact, their database has caught the attention of the federal government, which does not have comprehensive case information despite pressure from families and legal mandates to do so.<sup>11</sup> Federal agencies have asked Sovereign Bodies Institute multiple times for their data. Each time, the organization has consulted with families and the answer has been a resounding “no.” Lucchesi explained the rationale behind their refusal:

The federal government has never published any data or even just a number of how many cases have occurred in their jurisdiction. Members of the FBI and DOJ requested access to our database and we said we would be willing to sit at the table and discuss it if they (1) made data on cases in their jurisdiction available to us and to tribal governments and (2) increased the rate of prosecution of non-Indian sexual offenders in Indian Country (over 70% of the cases federal prosecutors decline in Indian Country are sexual assaults). The only response we ever received was ‘We understand.’ That tells us they are unwilling to provide data or hold perpetrators accountable—why would we trust our data with them if we can’t trust them to keep us safe and document our deaths once we have been killed in their jurisdiction?



Sovereign Bodies Institute has a clear analysis of the context and circumstances leading to missing data, as well as what it would take to start to build trust and share data between government officials and civil society groups. As their work and table 4.1 show, while oppression may be organized to produce missing data and flawed information across multiple domains, creative informatic resistance is organized to respond in a multiscalar, transversal way. Some informatic tactics of activists are designed to respond to failures of the state and some to failures of the media and some to challenges of missing data in the familial domain.

This multiscalar resistance across domains of the information ecosystem demonstrates that femicide data activists and journalists have a deep understanding of the legal, administrative, narrative, and interpersonal factors that produce missing data in their context and have also crafted creative responses to research information in the midst of those challenges. While prior research on data activism has tended to focus on how activist data circulate publicly and what political outcomes they achieve, examining the activist labor of data production demonstrates just how much expertise activists themselves must build about the unequal and deeply flawed conditions of information production about femicide.

One important throughline across activist research tactics in all domains is not only their use of existing information sources but their creative production of coordinated, collective human relationships as a mechanism for case discovery, information sharing, and deliberative decision-making about cases. This resonates with data feminism's principle to *embrace pluralism*—to bring together multiple perspectives with priority given to local, collective, and Indigenous ways of knowing. This is embodied by Mumalá's federated media monitoring structure, where different individuals are responsible for monitoring different Argentine provinces. Or in their collaborative method for deliberating, through WhatsApp, about which cases should be included in the database. It is also manifest in their relations with feminist journalists whom they can call on to try to get information from prosecutor offices. We can also see this evidenced in Sovereign Bodies Institute's work to build relationships with families to help them access services, and with grassroots Indigenous groups across North America, who then notify each other about new cases. There is collective power in the development of these novel human-relational-informational configurations.



FIGURE 4.3

MMIWG2 Jurisdiction Flowchart from the Sovereign Bodies Institute's MMIWG2 & MMIP Organizing Toolkit. Courtesy of Sovereign Bodies Institute.

## THE EMOTIONAL LABOR OF RESEARCHING VIOLENCE

As I mentioned at the beginning of this chapter, a persistent theme across the researching phase is the emotional labor required to research cases of violence as well as the mental and emotional toll of doing the work. Lara Andres, from *Ahora que sí nos ven* (Now that they do see us), an observatory in Argentina, says it like this: “The truth is that we can’t spend all day, sitting down, reading about feminicides, because it does your head in. It makes you sick, so I think the most time I’ve spent registering feminicides, or reading news, would be two or three hours a day.” Individuals talked about the psychological and emotional effects of the work on their well-being. Carmen Castelló found that watching the news to find cases early in the morning would leave her in a state of anxiety for the rest of the day, and Nerea Novo from *Feminicidio.net* discussed stories that she had read that she couldn’t stop thinking about. It can be especially challenging for activists who work directly with families. As Brandy Stanovich from the Native Women’s Association of Canada described: “When you’re in certain settings, you have to be the strong one in that room because you’re with the survivors of people who have gone missing or been murdered. But, later on, you might just want to lay down and cry because you can feel the emotions.” Some people find that it is not only about the graphic details of some cases, but also that the counting and aggregating becomes too much to bear. Paola Maldonado Tobar, from the *Alianza* in Ecuador, reflected, “This work of marking how many women have been victims of femicide and placing those on a map is painful, it’s terrible.” And for Audrey Mugeni, from *Femicide Count Kenya*, seeing the aggregated numbers is tremendously difficult: “When you go to the Excel sheet you’re like, oh no, I need to close my eyes. So it’s so much harder to look at the Excel sheet than it is to just read the stories.”

Given the range of emotional and affective responses that activists experience as they research and record each case of femicide, it is curious that the output of this labor—typically the neatly arranged spreadsheet or database—barely reflects this turbulence. In April 2022, I had a conversation with Helena Suárez Val, colead of the *Data Against Femicide* project, about how challenging it is to *make labor visible* in relation to the research work that goes into femicide data production. This is one of the principles of data feminism, drawing from feminist thinkers who describe how women’s labor—including emotional labor and care work—is frequently invisibilized and devalued. An antidote, then, should be to show and credit that labor. Yet in the case of femicide research, there are political and protective reasons to keep that labor hidden.

When you visit the publicly available spreadsheet that Helena produces for *Feminicidio Uruguay*, it is straightforward. Almost painfully simple to read. Each row

	A	B	C	D	E	F	G	H
	Nombre de la mujer	Edad de la mujer	Fecha del feminicidio	Lugar del feminicidio	Descripción	Relación del feminicida con la mujer	Edad del feminicida	Espacio donde se encontró el cuerpo
2		7	8 oct. 2022	Barrio Albisó, Municipio de San Antonio, Salto	Un hombre asesinó a su hija y luego se suicidó. La madre de la hija se había separado de él hacía dos semanas y había denuncias previas.	Padre	43	Domicilio del feminicida
3		35	4 oct. 2022	Capurro, San José	La encontraron en su casa, semidesnuda y con golpes en la cabeza con un bloque. Detuvieron a un hombre que anteriormente había alquilado un cuarto en la casa de ella.	Allegado		36 Domicilio de la mujer
4		33	13 ago. 2022	Guichón, Paysandú	Estaba desaparecida desde el 13 de agosto y encontraron su cuerpo el 18 de setiembre, bajo unos escombros en un basural. Fue detenido un hombre con quien tenía una relación.	Pareja o ex-pareja		40 Espacio público
5		50	14 set. 2022	Puente Valdez Chico, San José	Encontraron su cuerpo bajo un puente con heridas, y el auto de su pareja se halló abandonado. Habían ido juntos al balneario Kiyú y su familia la denunció desaparecida cuando no recibieron noticias por varios días. El hombre apareció muerto unos días después en el mismo río.	Pareja o ex-pareja		Espacio público
6		38	6 set. 2022	Barrio Talar, Pando, Canelones	La encontraron apuñalada afuera de su casa y antes de morir acusó a su pareja de haberla matado. El Colectivo Trans del Uruguay denuncia que la prensa reportó la noticia sin respetar su identidad de género.	Pareja o ex-pareja		40 Domicilio de la mujer
7		28	23 ago. 2022	(Las Primicias y camino Capitán Corralio Lacost)	La encontraron en su casa junto a su hijo herido, que luego murió en el hospital. Su pareja les disparó y luego se suicidó.	Pareja o ex-pareja		32 Domicilio de la pareja
8		8	23 ago. 2022	(Las Primicias y camino Capitán Corralio Lacost)	Lo encontraron herido en su casa junto a su madre y murió en el hospital. Su padre les disparó y luego se suicidó.	Padre		32 Domicilio de la pareja
9		40	19 ago. 2022	Agua Salto, Salto, Salto	Encontraron su cuerpo con signos de violencia en un camino vecinal. Estaba embarazada de siete meses.	Pareja o ex-pareja		56 Vía pública
10		31	4 ago. 2022	(Las Violetas y María Orticochea) Paso de la Du	La mató su ex-pareja en la calle frente a sus dos hijas. Él tenía orden de alejamiento y la mató cuando ella iba a denunciar que estaba cerca de la casa. El feminicida intentó suicidarse y fue hospitalizado, muriendo días después.	Pareja o ex-pareja		40 Vía pública
11		26	23 jul. 2022	Colonia Nicolich, Canelones	La encontraron en su casa junto a su hija, habían sido golpeadas y apuñaladas.	Pareja o ex-pareja		25 Domicilio de la mujer

FIGURE 4.4  
The open spreadsheet for Feminicidio Uruguay. Courtesy of Helena Suárez Val/Feminicidio Uruguay.

corresponds to a woman, to her violent death, and to related details surrounding the event: name, age, date, place, relation of the perpetrator to the woman. There is a matter-of-fact narrative in Column E that describes how her body was found, method of death, or whether she had been missing. Yet each row in this spreadsheet is the result of many hours of reading and research, triangulating details from multiple media articles, updating as new information surfaces about an investigation. Furthermore, because Helena produces a map based on her database, she often spends still more hours using photos, Google Street View, and media articles to find the most precise place to geolocate a particular case.

What is not visible in the spreadsheet is how Helena feels when she walks by one of the places of violence that she has geolocated on her map. One of them is about half

a block from her apartment in Montevideo, and she remembers the blood on the sidewalk every time she walks by. What is also not visible in the spreadsheet is her state of agitation when a case is initially reported in the press until the time that a perpetrator has been identified: “I would tell you that the day the alert arrives, I am involved in some way that whole day and probably a couple of days more to come. . . . You kind of stay a little tense until there’s some sort of resolution.” Still, for some cases, locating the perpetrator may take much longer or may never happen. What is also not visible in the spreadsheet are the many hours of reading stories of brutal violence; the careful collection of each painstaking detail of what is known about a case; the way it feels to leave so many fields blank.

In both her capacity as an activist and as a scholar, Helena has been thinking deeply about data production work as care work. We wrote a short essay together about this following Maria Puig de la Bellacasa’s formulation of care work as an ethical-political commitment to “neglected things.” For activists, painstakingly researching cases is one small way to recuperate and repair such public neglect. Matters of care, for Puig de la Bellacasa, involve not only revaluing the often invisibilized (and gendered) labor of maintenance and repair, but also disrupting the Western academic tendency to value distance, neutrality, objectivity, and impartiality between researcher and research subject. Instead, care work weaves researcher and researched together into affective relationships that produce new knowledge through proximity, emotion, and connection.<sup>12</sup> This resonates with the data feminism principle to *elevate emotion and embodiment*. Femicide counterdata producers are deeply intimate with their data points—the people and places whom the data represent—because they have spent so much time researching and triangulating sources, rejecting stigmatizing media narratives, finding photos, and correcting the errors of the state and the media. As a result, many activists are deeply protective of their data, as Lucchesi and Sovereign Bodies Institute evidenced in their refusal to share it with federal authorities. For activists, paradoxically, these data are not data, never were data, and cannot be reduced to data. Yet their representation as data persists in the form of rows in databases or grids and columns of spreadsheets. The gridded orderliness hides the labor of its production.

But is this a feature or a bug? Do activists want their spreadsheets to be messy records of emotion, to be testimonies to their rage, or their attachment or their sadness? I will have to assume not, since nobody that we interviewed documents femicide in such a way. Rather, the erasure of the material and emotional labor behind femicide data production feels more strategic and calculated. It functions as a kind of *hack* to hegemonic data production. Presenting the results of a fraught and intimate and labor-intensive research process in the cool logic of the spreadsheet grid appropriates

Western, colonial, patriarchal penchants for distance and quantification and deploys them to elevate and amplify the feminist concerns and the feminist rage seething just under the surface of the lined, light-gray boxes.<sup>13</sup> But it elevates and amplifies rage precisely by visually obscuring it. This builds on and extends the data feminism principles *make labor visible* and *elevate emotion and embodiment*. It may not be desirable, in this case, to make labor visible for everyone, all the time. Sometimes, it is necessary to hide that labor behind the strategic deployment of the rhetoric of objectivity. Those within the feminicide data community share their labor and their grief with each other, but from the outside the work appears as neatly ordered rows and columns. A paper by communications scholars Yu Sun and Siyuan Yin on feminist data activism in China makes a related point, which is that under nondemocratic regimes, it is problematic and even dangerous to publicly credit activist labor because it makes activists targets for political oppression.<sup>14</sup> Likewise, it may not always be politically productive to elevate emotion when such emotion becomes the basis for discrimination—the familiar patriarchal approach of dismissing emotion as counter to reason or painting feminist concerns as too “personal.” Moreover, some parties do not deserve our grief and our outrage, which are constantly leveraged in exploitative and voyeuristic ways by the media, leading to more than one activist characterizing media coverage of feminicide as “trauma porn.” This leads me to ask the question: In the process of doing data science, when should one seek to make labor visible? In which contexts do we elevate emotion? And for whom?

#### LESSONS FOR A RESTORATIVE/TRANSFORMATIVE DATA SCIENCE

The information discovery tactics described in this chapter surface some key lessons for a restorative/transformational data science, especially if we look at differences between the researching labor of feminicide data activists in comparison with how researching and compiling data is treated in hegemonic data science projects. First, the labor of hegemonic data science is masculinized, overvalued, and overcompensated. The research group AI Now reported in 2018 that women comprise only 15 percent of AI research staff at Facebook and 10 percent at Google.<sup>15</sup> More broadly speaking, men dominate computer science occupations in the United States, comprising almost three-quarters of the workforce, and gender balance in computing has been declining since the mid-1980s.<sup>16</sup> In contrast, the work to research and monitor feminicide is almost the polar opposite. As we have seen, it is done almost exclusively by women and gender nonconforming people. Data activism that centers on racialized feminicide is almost exclusively undertaken by Black women, Indigenous women, and women

of color. Most research is unwaged work. When projects originate from a nonprofit or from journalism, people may be paid for their time, but organizations have a hard time securing operational funds to sustain it. Thus, the labor of femicide counterdata research is feminized, racialized, devalued, and undercompensated. This is consistent with the feminist concept of reproductive labor—the care work that sustains, maintains, and reproduces society.<sup>17</sup> In this case, instead of cleaning homes and raising children, activists are stewarding the memories of killed women, caring for wounded communities, and working toward repair and justice. It is women and queer people who are doing the reproductive labor to clean up, informatically speaking, after the structural excesses and negligence that led to the gender-related violence in the first place.

But this stark divergence is not only related to femicide data research. Not all parts of the conventional data science pipeline are masculinized and highly compensated. Melanie Feinberg, an information scientist, describes how data collection and classification work is perceived as “unskilled and mechanical.” When she runs data collection assignments in her classes, the students imagine (erroneously!) that “there is nothing to be learned from the process of generating data, because data collection is the mere recording of objects speaking for themselves.”<sup>18</sup> Indeed, data collection, annotation, sorting, and labeling is often outsourced from the Global North to the Global South, where it is done for low wages, disproportionately by women and racialized people.<sup>19</sup> This goes even and especially for information tasks that resemble the work of femicide data research in their emotional burden: content moderation on commercial platforms such as Facebook and YouTube. In these jobs, low-paid workers spend hours per day sifting through graphic, violent, racist, misogynist, exploitative content and labeling it according to platform policies.<sup>20</sup> This has led information scholar Julian Posada, among others, to assert that the AI industry profits from political instability and colonizes catastrophe to provide bad jobs to economically vulnerable people.<sup>21</sup> In contrast, the high-status part of data science supposedly comes *after* the data production activities of collection, classification, acquisition, and preprocessing. Drawing from these perceptions of status, Nithya Sambasivan and colleagues wrote a paper titled “Everyone Wants to Do the Model Work, Not the Data Work,” outlining the downstream harms of disinvesting in data quality.<sup>22</sup>

Femicide counterdata activists provide a compelling model for what it looks like to refuse this (gendered, racialized, colonial) stratification of data labor. For better or for worse, they are intimately connected to their data due to the hours of time invested in researching each case. This labor engenders a deep expertise in the information ecosystem surrounding femicide, as evidenced by activists’ analysis of sources of missing



data in table 4.1. Not only do activists understand the flaws and limitations of the information ecosystem, they understand how their own data also inherit those limitations. This is the essence of the *consider context* data feminism principle. In recent years, there has been more work that examines contextual knowledge as a unique form of expertise that has been undervalued by the mainstream data science community. For example, Annabel Rothschild and colleagues demonstrate how civic data workers, due to their proximity to the data collection process and their deeper domain knowledge, engage in information contextualization practices that mainstream data scientists stand to learn a great deal from.<sup>23</sup> In the case of feminicide data activists, knowledge of the limitations of their data lead to caution in how they use them and communicate them downstream. For example, all groups mentioned that they know their databases are not complete since not all cases are reported in the media or publicized on social media. They take care to communicate that any statistics remain undercounted, with the most intersectionally marginalized populations as the ones who remain most undercounted.

As Lucchesi reported, “The data requires a really intimate relationship in order to make it workable.” She scoffed at some of the data requests that she has received from outsiders: “The requests that we get are based on assumptions that the data is just kind of like this divining tool that anybody can just jump in and use and that somehow all of the mysteries of this crisis will be solved from their armchair with casual exploration. And that’s not the case. If it were, we would have fixed it already.” The ignorance that Lucchesi is challenging here is a direct product of hegemonic data science’s undervaluing of data research and collection work—leading data scientists to believe that they could just explore “the data” (Which data? Well, whichever are available and cheap.), treat them as ground truth, make a model, and then *voilà!* hidden wizardly insights about MMIWG2 are revealed.<sup>24</sup> At some level, this also points to a real failure in the education system: that educators in computer science, statistics, and mathematical domains have not been able to understand the urgency of integrating humanities and social sciences concepts, leaving their students impoverished—overconfident and underprepared—to design any kind of information system that relates to human life.

Finally, as I previously mentioned, many of the most creative tactics that counterdata groups develop to source information about feminicide come not through novel sensors or computational techniques but via the production of novel and collective forms of human relationships. These relationships tend to be nonextractive and authentic; they are not about acquiring data and then generating capital (financial or social) from it. This represents a contrast with hegemonic data science that, to the extent that it develops novel forms of human relationships, tends to be in the model of what Paola Ricaurte calls *data extractivism*, which operationalizes everything as a potential data

source.<sup>25</sup> Driven by profit, modeled on colonial relations, hegemonic data science treats humans as interchangeable units of supply and demand (e.g., Uber drivers and their customers), as mechanical automatons (e.g., Amazon Mechanical Turk and microwork platforms), or as chauffeurs for sensors (e.g., Waze, Google Maps). This is not to say that feminicide groups' coalitions and human relationships are free of conflict (they are definitely not), but rather that the commitment to nonextractive human relations represents a key aspect of their data epistemology. It is a form of epistemic disobedience and a form of resistance to the data extractivist regime.

This epistemic challenge is resonant with non-Western calls for emphasizing relationality and collective responsibility in data and artificial intelligence. These include decolonial AI, Indigenous data sovereignty, and other data epistemologies described in chapter 8. Ultimately, a restorative/transformational data science requires such an alternative epistemology, and one that places care, context, and connection at the center of the data gathering process. This presents a profound challenge to the data acquisition process of hegemonic data science. As Feinberg states, "The real revolution in data labor will be in acknowledging that data collection should be celebrated for its skill and creativity."<sup>26</sup> It's not only about revaluing data research and production because it's the right thing to do, but also because it challenges harmful data extractivist regimes and it results in better data science. For example, feminicide data activists are profoundly aware of their data's limitations and quality issues, whereas hegemonic data scientists are often stunningly ignorant of or surprised by theirs.<sup>27</sup> This is especially true when the information ecosystem surrounding an issue is highly influenced by structural inequality, such as when there is rampant missing data, biased data, harmful and stereotypical information, or mis- and disinformation, all of which need to be assessed and evaluated for inclusion in a broader dataset. Indeed, a restorative/transformational data science involves dismantling the existing stratification of data labor, revaluing the proximity of the data producers to their subject matter, and doubling down on contextual knowledge as a unique and unautomatable form of expertise, something to teach and cultivate in community.

## CONCLUSION

*Researching* is the second stage of a restorative/transformational data science project in which an individual or group seeks information about individual cases to add to their database. This can include sourcing existing datasets, mining state and media sources of information, verifying details, and triangulating across sources. In the researching stage, activists conjoin their analysis of power with their information-seeking practices,

leading them towards a skillful, on-the-ground understanding of the sources of missing data and biased information that permeate the femicide information ecosystem. Consequently, they navigate informatic gaps, biases, and errors by employing a variety of creative strategies to source information from the state, the media, and families. Activists often develop novel human-relational-informational configurations—using relationships of trust and solidarity to establish networks of information providers. Still, the work demands extensive emotional labor to sift through stories of violence and precisely document details of a human life and death.

Researching femicide cases goes hand-in-hand with recording such details into spreadsheets and databases, the subject of the next chapter. Thus, while these stages are described separately, in practice they are tightly linked, with activists going back and forth between seeking information about a case and copying new details into structured fields and categories. The work is not only challenging because of the violence but also because there is no end in sight. Activists seek a world in which gender-related violence has been eliminated but women continue to be killed and cases continue to surface. In the face of state injustice and media stigmatization, researching cases is an informatic strategy to challenge such structural bias. Activist researching practices point toward a restorative/transformational data science that values proximate relations between data scientists, datasets and data subjects, as well as elevates the labor of data sourcing, labeling, and curation practices.



This is a section of [doi:10.7551/mitpress/14671.001.0001](https://doi.org/10.7551/mitpress/14671.001.0001)

# Counting Femicide

## Data Feminism in Action

By: Catherine D'Ignazio

### Citation:

*Counting Femicide: Data Feminism in Action*

By: Catherine D'Ignazio

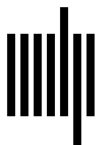
DOI: 10.7551/mitpress/14671.001.0001

ISBN (electronic): 9780262378017

Publisher: The MIT Press

Published: 2024

The open access edition of this book was made possible by generous funding and support from MIT Libraries, and MIT Press Direct to Open



The MIT Press

© 2024 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license.

This license applies only to the work in full and not to any components included with permission. Subject to such license, all rights are reserved. No part of this book may be used to train artificial intelligence systems without permission in writing from the MIT Press.



The open access edition of this book was made possible by generous funding from the MIT Libraries.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in ITC Stone Serif Std and ITC Stone Sans Std by New Best-set Typesetters Ltd.

Library of Congress Cataloging-in-Publication Data

Names: D'Ignazio, Catherine, author.

Title: Counting feminicide : data feminism in action / Catherine D'Ignazio.

Description: Cambridge, Massachusetts : The MIT Press, [2024] | Includes bibliographical references and index.

Identifiers: LCCN 2023027064 (print) | LCCN 2023027065 (ebook) | ISBN 9780262048873 (hardcover) | ISBN 9780262378000 (epub) | ISBN 9780262378017 (pdf)

Subjects: LCSH: Feminism. | Feminism and science. | Big data—Social aspects.

Classification: LCC HQ1190 .D5735 2024 (print) | LCC HQ1190 (ebook) | DDC 305.42—dc23/eng/20230828

LC record available at <https://lcn.loc.gov/2023027064>

LC ebook record available at <https://lcn.loc.gov/2023027065>

10 9 8 7 6 5 4 3 2 1