

## 8 A TOOLKIT FOR RESTORATIVE/ TRANSFORMATIVE DATA SCIENCE

This final chapter is a toolkit for anyone who may be contemplating starting or joining a restorative/transformational data science project. Perhaps you are a data activist or a journalist, or maybe you are an urban planner or a librarian, a digital humanities scholar or an academic, the head of a nonprofit or a member of a community group or a designer. Or somebody else entirely. My hope is that these pages will offer both reflection and practical guidance.

As I sat down to map the contours of this toolkit, I realized again and again that all of the ideas I thought I had were actually things that activists had told me or showed me or taught me. One of the most humbling and inspiring aspects of writing this whole book has been the continual realization of the depth and the breadth of what our partners have to teach the rest of us who aspire toward transformational social change. Feminicide data activists center care, memory, and justice without sacrificing rigor. They don't imagine data as a "solution" but rather see data as one tactic in a larger, networked movement of social and political actors. They stay close to the communities and families impacted by feminicide, in some cases providing direct healing and support. They are highly creative in acquiring information from a deeply biased and unjust information ecosystem. They develop ways to circulate feminicide data for diverse impacts, ranging from policy reform to narrative change to mass mobilizations.

This toolkit is a first step toward drawing out some of these lessons and speculating on how these may be useful to others using data in the service of restoration and transformation, healing and liberation. While much research produced in the Global North about the Global South ends up inaccessible behind paywalls, this toolkit, which is included in a book published in a free, open-access format, is an attempt to counteract that dynamic and contribute knowledge back to grassroots communities, activists, journalists, academics, and others.

The questions and activities in the toolkit are drawn from themes that surfaced in our work on feminicide, but they are purposely written to be broadly applicable. My hope is that these examples, questions, and activities will be helpful to anyone using data-driven methods for monitoring, auditing, and inquiry, and especially projects related to structural inequality. I am confident that at least some of the many lessons learned from counting feminicide will be helpful for other efforts working to restore rights and to transform systems. These might include projects that document forms of violence other than feminicide—for example, the already existing efforts to document police killings of Black Americans, LGBTQIA+ hate crimes, or the murders of Indigenous land defenders. Yet there are also counterdata efforts occurring that are related not only to physical violence but to other forms of economic violence or structural inequality. Recently there has been a remarkable growth, for example, in nonprofit and activist groups that count and monitor evictions in the United States. Here the rows of data represent cases of eviction rather than cases of feminicide, but the intent is similar: first, to use quantification to provide direct services and legal aid to impacted people. And second, to transform policies and systems that treat housing as a commodity rather than as a human right. Here there is great resonance with the feminicide data activists' examination of power and use of data to reframe and remake personal problems into structural patterns. To inspire connections to work in other domains, I have included a list of restorative/transformational data science examples in the toolkit.

That said, because the toolkit arose in the context of feminicide data activism, it is highly likely that not all activities in the toolkit apply to *all* restorative/transformational data science projects in *all* domains. I encourage people to use what is useful and, if you are moved to do so, contribute comments, questions, and critiques to the evolving open toolkit located at <https://mitpressonpubpub.mitpress.mit.edu/pub/restorative-data-toolkit>.

## GLOSSARY

**Coliberation:** This is the idea that all of us are harmed by systems of unequal *power* and that we, working together, can free ourselves of its multiple burdens—material, psychic, spiritual, and intergenerational. There is a well-known quote from Aboriginal activists in Queensland, Australia, that best represents this idea: “If you have come here to help me you are wasting your time, but if you have come because your liberation is bound up with mine, then let us work together.”<sup>1</sup>

**Counterdata:** Data that are produced by civil society groups or individuals in order to challenge unequal *power*. Producing counterdata is not (only) about countering

*missing data* or inadequate *official data* from institutions but is also used to challenge state bias and inaction, to galvanize media and public attention, to reframe political debates, to work toward policy change, and to help heal wounded communities. Counterdata production, like *data activism* more broadly, is a citizenship practice. It is an informatic form of enacting democratic dissent, prompting protest and insisting on political engagement. But not all activist data is counterdata: Indigenous scholars and data activists emphasize that their work is about sovereignty and not about countering the settler state.

**Data activism:** The use of data and software to pursue collective action and exercise political agency. Producing *counterdata* and engaging in *restorative/transformative data science* are specific ways—but far from the only ways—of engaging in data activism.

**Data epistemologies:** Theories and approaches to knowing things about the world with data. Mainstream data epistemologies are heavily positivist—seeking to use data to find universal truths over a consideration of context. This results in *hegemonic data science*. Many scholars and activists have highlighted how mainstream data epistemologies replicate violent and extractive and colonial modes of knowledge generation. Emerging alternative data epistemologies include data feminism, feminist data refusal, emancipatory data science, environmental data justice, decolonial AI, Indigenous data sovereignty, and queer data (see the next section for a short guide).

**Discordant data:** An idea from Helena Suárez Val that describes the fact that *official data* and *counterdata* often deliberately do not coincide; they are *discordant* because they intentionally use different definitions, measurements, and classification strategies.<sup>2</sup>

**Emotional labor:** Producing and working with data related to social inequality almost always involves being witness to trauma and violence, whether the data are about femicide or evictions or environmental harms. The psychic and emotional burdens of this witnessing work should not be overlooked, particularly for survivors who may have firsthand experience of such violence. Projects engaging in *restorative/transformative data science* can develop strategies for self- and team care and/or they may choose to “flip the script” and take an assets-based approach to producing data, where the focus is on mapping communities’ strengths and joys, not their deficits and traumas.

**Hegemonic data science:** Mainstream data science that works to concentrate wealth and *power*; to accelerate *racial capitalism*, perpetuate *patriarchy*, and sustain *settler colonialism*; and to exacerbate environmental excesses and social inequality.

**Information ecosystem:** A dynamic constellation of actors that includes infrastructure, tools, technology, producers, consumers, curators, and sharers of information about a particular topic. 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.

**Intersectionality:** The idea from Black feminism that systems of *power* compound and combine and cannot be understood in isolation. For example, a single-axis analysis that looks only at *patriarchy* will miss the ways in which *patriarchy* intersects with *white supremacy*, with *settler colonialism*, with *ableism*, and so on. Intersectionality comes from theorizing the experiences of Black women in the United States, and substantial contributions to it have been made by the Combahee River Collective, Kimberlé Crenshaw, and Patricia Hill Collins, among many others.

**Memory work:** Practitioners of *restorative/transformational data science* understand that producing and circulating *counterdata* is a way to assemble and care for stories from the past in order to develop new visions for the present and future.

**Missing data:** Data that are neglected by institutions, despite political demands that such data *should be* collected and made available. Missing data may include data that are entirely absent but also data that are sparse, neglected, poorly collected and maintained, purposefully removed, difficult to access, infrequently updated, contested, and/or underreported.

**Official data:** Data that are produced by the state, international governing bodies, and/or other mainstream institutions such as large corporations or professional associations.

**Power:** The current configuration of structural privilege and structural oppression, in which some groups experience unearned advantages—because various systems have been designed by people like them and work for people like them—and other groups experience systematic and violent disadvantages—because those same systems were not designed by them or with people like them in mind.<sup>3</sup> Specific systems of privilege and oppression include but are not limited to the following:

- **Ableism:** The systemic privileging of ability that results in the oppression of disabled people based upon real or perceived impairments. It “others” disabilities, chronic illnesses, and neurological or mental illness.<sup>4</sup>
- **Cisheteropatriarchy:** (Synonymous with *patriarchy* in this book.) The social and political system that elevates cisgender, heterosexual men and oppresses those with

minoritized gender identities (women, trans, travesti, nonbinary, two-spirit people, and more) and minoritized sexual orientations (lesbian, gay, bisexual, and more).

- **Colonialism:** Refers to some combination of territorial, cultural, linguistic, political, and/or economic invasion and subsequent domination of one group of people by another group of people.<sup>5</sup>
- **Economic violence:** Economic policies and practices that systematically deprive groups of people of their human rights to life, food, clothing, housing, and medical care. This might be through exclusion from labor markets, underwaged or unwaged work, exclusion from education, privatization of public goods, and more. *Neoliberal* policies result in economic violence, and so do *colonialism*, *patriarchy*, and the other systems of *power* referenced here.
- **Extractivism:** Positions land, air, water—and increasingly also data and digital information—as free resources to be mined, wasted, privatized, and profited from.<sup>6</sup> Often results in *economic violence*.
- **Neoliberalism:** Refers to economic policy that favors free-market capitalism and tends to oppose government regulation that intervenes in markets. Approaches include privatization of public goods, deregulation, and restriction of public spending. Neoliberalism strengthens corporate power, extracts private profit from public resources, externalizes private costs (such as environmental harms) to the public sector, removes social protections for the most vulnerable, and undermines representative democracy by creating a vicious cycle of super rich oligarchs who manipulate the political machinery, doubling down on neoliberal policy to get super-richer. Neoliberalism produces *economic violence*.
- **Patriarchy:** See *cisheteropatriarchy*.
- **Racial capitalism:** The idea from Cedric Robinson that capitalism is founded upon racial stratification. In other words, capitalism and racism are inextricable: free markets are founded upon unfree labor that arises from slavery, colonization, *economic violence*, and *extractivism*. This system is historic and ongoing.
- **Settler colonialism:** A form of historic and ongoing colonization in which outsiders come to land/air/water/subterranean earth inhabited by Indigenous peoples, dispossess them of that land, and then insist—through institutions, laws, culture, and violence—that settlers are sovereign over the stolen land.<sup>7</sup>
- **White supremacy:** A historically based, institutionally perpetrated system of exploitation of continents, nations, and peoples of color by white peoples and nations of the European continent, for the purpose of maintaining and defending a system of wealth, power and privilege.<sup>8</sup>

**Restorative/transformational data science:** An approach to working with systematic information that seeks, first, to heal communities from the violence and trauma produced by structural inequality and, second, to envision and work toward a world in which such violence has been eliminated. *Restoration* involves the use of data for restoring life, living, and vitality to the individuals, families, communities, and larger publics harmed by unequal systems of *power*. It also seeks the restoration of rights—the right to live a life free from violence, for example, or the right to adequate housing, or the right to ancestral homelands. *Transformation* involves the use of data to dismantle and shift the structural conditions that produced the violence in the first place. It is both visionary and preventative.

**Triangulation:** Using multiple sources of information about the same event or phenomenon to cross-reference and verify details. This is often necessary when *official data* are suspect or sparse or when there is not a single authoritative source of information in the *information ecosystem*.

#### WHAT IS YOUR DATA EPISTEMOLOGY?

Practitioners of restorative/transformational data science mobilize alternate epistemologies of data that challenge the extractive and violent regimes of hegemonic data science. In the last decade, the number of alternate data epistemologies has multiplied. Each one offers at the very least some foundational texts and principles, or else some flagship projects that can serve as models to follow. Some also offer supportive communities, and others, like the global movement for Indigenous data sovereignty, have robust theoretical foundations, thriving communities, and policy and governance frameworks.

Many of these data epistemologies come from theorizing directly from the experiences and histories of specific groups and the data harms that they have experienced. For example, Data for Black Lives aims to use data “to create concrete and measurable change in the lives of Black people.” Data epistemologies are not interchangeable ethical checklists: moving forward with a particular epistemology carries responsibilities for who your work serves, who you are committed to being in dialogue with, and who you are accountable to. If you seek to mobilize an epistemology that comes from theorizing experiences that you have never had (e.g., you are in a cisgender heterosexual team who wishes to draw from a queer data perspective) then you will need to think carefully about how you can do that without causing harm and without engaging in appropriation. In some cases, the best answer may be to step aside and make space for

a group who does have that lived experience to lead the project and your team can get them coffee. That's a joke—but only a half-joke! The real issue is thinking carefully about how you can support and center leadership by people who have the knowledge and cultural grounding that comes from lived experience.

For example, one troubling thing I have witnessed recently is grants being given to settler people who propose to draw from Indigenous data sovereignty and work with Indigenous communities in a participatory way. These projects receive funding without having established partnerships with tribes or nations or Indigenous-led groups, so at no point in the project has an Indigenous person or community weighed in on the validity and utility of the idea. This sets up the structural conditions for harm to occur because once funding has been secured, the settler person is most accountable to the settler funder and has to retroactively find a community partner amenable to an idea created in a settler vacuum.

To seek an appropriate data epistemology for your project, you may want to reflect on your own positionality as well as the community that you want to serve and be accountable to with your data project. See the data epistemology activity in the Getting Started section ahead for a starting point for this kind of reflection. And ahead is a list of data epistemologies that showcases those approaches with which I am most familiar and have been participating in, reading about, or following online.

## DATA EPISTEMOLOGIES

### DATA FOR BLACK LIVES

A movement of activists, organizers, and mathematicians based mainly in the United States that aims to use data “to create concrete and measurable change in the lives of Black people.” It is organized into regional hubs. See in particular <https://d4bl.org/> and their 2021 report on *Data Capitalism and Algorithmic Racism*.

⇒ **Especially relevant for:** Projects by/with Black Americans, projects in solidarity with Black Lives Matter, projects seeking a regional community.

### DECOLONIAL APPROACHES

- ❖ **Tierra Común**—A global community of scholars and activists working to decolonize data. See <https://www.tierracomun.net/>, the scholarship of Paola Ricaurte, and the book *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*.
- ❖ **Ubuntu ethics for AI**—Scholars and technologists such as Sabelo Mhlambi and Serena Dokuaa Oduro have advanced the idea of incorporating Ubuntu philosophy

into AI systems and AI policy. See Mhlambi's paper *From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance*.

- ❖ **AI Decolonial Manifesto**—A collaborative statement by two dozen scholars across computer science, social sciences, humanities, and human rights, to challenge the “Western-centric biases” being baked into AI. See <https://manifesto.ai/>.

⇒ **Especially relevant for:** Projects focused on colonialism, imperialism, and dispossession, projects from the Global South, projects seeking South-South community, and projects mobilizing non-Western ethics.

### THE DESIGN JUSTICE NETWORK

A network of practitioners with an important set of principles grounded in intersectional feminism, which is focused on using design to support social justice. While the Design Justice Network is focused more on design, the principles have great relevance and applicability to undertaking data science projects in community. See <https://designjustice.org/> and the book *Design Justice* by Sasha Costanza-Chock.

⇒ **Especially relevant for:** Projects coming out of design disciplines, projects drawing from intersectionality, projects using participatory methods, and projects seeking to connect with a robust community of practice.

### EMANCIPATORY DATA SCIENCE

Developed by professor Thema Monroe-White, emancipatory data science draws from emancipation theory, critical race theory, and critical quantitative theory (see #QuantCrit ahead), to theorize how to use data science in the service of “uplift and empowerment.” Monroe-White outlines three functions emancipatory data science can undertake: diagnosis and critique (of data harms), viable futures (related to more equitable data practices and policies), and transformation (of the data science community itself). See the paper “Emancipatory Data Science: A Liberatory Framework for Mitigating Data Harms and Fostering Social Transformation.”

⇒ **Especially relevant for:** Abolitionist projects, projects drawing from critical race theory, and projects seeking to describe data harms for minoritized people.

### ENVIRONMENTAL DATA JUSTICE (EDJ)

An emerging approach that joins environmental justice (elaborated in the 1990s by scholar Robert Bullard and principles adopted at the First People of Color Environmental Leadership Summit) with data justice and critical data studies. Environmental data justice develops ways of working with environmental data, using participatory methods, in the service of community self-determination. To learn more, see the EDJ working group of the Environmental Data Governance Initiative (<https://environmental-data-justice.org/>), projects by the Technoscience Research



Unit (<https://technoscienceunit.org>), and the work of Lourdes Vera, Sara Wylie, and colleagues.

⇒ **Especially relevant for:** Projects using environmental data, projects drawing from environmental racism and environmental justice (EJ) principles, and people seeking to be in community with activist-scholars from a variety of disciplinary backgrounds.

## FEMINIST APPROACHES

Publications, programs, and emerging networks that mobilize feminist theory and activism to work with data and AI have been proliferating in recent years.

- ❖ **Data feminism**—Discussed in the next section.
- ❖ **The Design Justice Network**—Discussed earlier in this list. Grounded in intersectional feminist principles.
- ❖ **Feminist Data Manifest-NO**—A collaborative statement from scholars across numerous disciplines that “refuses harmful data regimes and commits to new data futures” (<https://www.manifestno.com>).
- ❖ **Data Feminism Network**—A learning community focused on equitable and feminist approaches to data. They run reading groups and produce events (<https://www.datafeminismnetwork.org>).
- ❖ Spanish language networks include Feminismo de Datos (<https://www.facebook.com/FeminismoDeDatos>), La Red Mexicana de Feminismo de Datos (<https://www.redmexicanadefeminismodedatos.org>) and La Red Latinoamericana de Feminismo de Datos (in formation, see <https://www.datagenero.org/>).
- ❖ **<A+> Alliance**—A global, multidisciplinary, feminist coalition of academics, activists, and technologists working to use artificial intelligence and IT to accelerate gender equality (<https://aplusalliance.org>).
- ❖ **Data Feminism Program**—By Data-Pop Alliance. Draws from data feminism to undertake international development projects ([https://datapopalliance.org/program\\_data\\_feminism](https://datapopalliance.org/program_data_feminism)).

⇒ **Especially relevant for:** Projects focused on cisheteropatriarchy, projects focused on gender and intersectionality, projects by/with women, queer, and/or trans people, projects using participatory methods, and projects that use qualitative and creative methods with data (art, storytelling, etc).

## INDIGENOUS DATA SOVEREIGNTY

If you are looking to put data in service of Indigenous nations and people, the global Indigenous data sovereignty movement has produced excellent scholarship, practical guidelines, and policy:

- ❖ Foundational texts include *Indigenous Statistics* and the edited volume *Indigenous Data Sovereignty: Toward an Agenda*.
  - ❖ Many Indigenous data sovereignty regional and global networks exist to link practitioners and scholars. For a list, see <https://indigenoustalab.org/networks/>.
  - ❖ The First Nations Information Governance Center developed the ownership, control, access, and possession (OCAP) principles. See <https://fnigc.ca/ocap-training/>.
  - ❖ Indigenous data sovereignty scholars have added the collective benefit, authority to control, responsibility, and ethics (CARE) principles to the FAIR principles for open science.
  - ❖ Indigenous data sovereignty research labs include the Data Warriors Lab and the Collaboratory for Indigenous Data Governance.
- ⇒ **Especially relevant for:** Projects by/with Indigenous people, projects focused on sovereignty, projects about dispossession and land rights, projects drawing from Indigenous worldviews, and projects that use oral history and storytelling methods.

#### #QUANTCRIT

Also known as quantitative criticalism and quantitative critical race theory, this is an approach developed in education, ethnic studies, and the social sciences that aspires to unite critical race theory with quantitative methods. Principles include the centrality of racism and oppression, the acknowledgment that numbers are not neutral, the idea that community voice is central, and the affirmation that statistical analysis can play a role in struggles for social justice. For a starting point, see the handy resource guide assembled by chemistry professor and educator Paulette Vincent-Ruz, available at <https://sites.lsa.umich.edu/pvincentruz/quantcrit-resources/>.

⇒ **Especially relevant for:** Projects from education, social work, and the social sciences, projects using quantitative and statistical methods, projects drawing from critical race theory, and projects seeking a robust academic conversation.

#### QUEER DATA

Recent publications aim to show how data may be used (or refused) to center the lives and well-being of queer, trans, nonbinary, and LGBTQ+ people. See the essay “Counting the Countless” by Os Keyes on radical data science, the book *Queer Data* by Kevin Guyan, and the volume titled *Queer Data Studies*, edited by Patrick Keilty. And make sure to read Dean Spade and Rori Rohlf’s powerful critique of “gay numbers” and the ways in which counting the LGBTQ+ population can perpetuate white supremacy.

⇒ **Especially relevant for:** Projects by/with LGBTQ+ people, projects focused on gender and sexuality, projects engaging theories and practices of refusal, and projects engaging queer theory.

## PRINCIPLES OF DATA FEMINISM

Data feminism is one epistemology of data, and the one that comprises the conceptual backbone of this book. In *Data Feminism*, Lauren Klein and I outlined what a feminist approach to data science might look like—an alternate data epistemology to challenge the standard operating procedures of hegemonic data science. We draw from intersectional feminist theory, activism, and writing to outline seven principles for working with data in a feminist way. I offer these principles here in the hopes that they will be a useful part of your toolkit, just as they have guided my own research and writing throughout this book. But I offer them with the acknowledgment that these are far from the only principles that you could use in your own data-driven work. The frameworks emerging from other data epistemologies might be more relevant and useful, depending on who you are, where you are located (geographically, socially, and spiritually), and what your project is about.

1. **Examine power.** Data feminism begins by analyzing how power operates in the world.
2. **Challenge power.** Data feminism commits to challenging unequal power structures and working toward justice.
3. **Elevate emotion and embodiment.** Data feminism teaches us to value multiple forms of knowledge, including the knowledge that comes from people as living, feeling bodies in the world.
4. **Rethink binaries and hierarchies.** Data feminism requires us to challenge the gender binary, along with other systems of counting and classification that perpetuate oppression.
5. **Embrace pluralism.** Data feminism insists that the most complete knowledge comes from synthesizing multiple perspectives, with priority given to local, Indigenous, and experiential ways of knowing.
6. **Consider context.** Data feminism asserts that data are not neutral or objective. They are the products of unequal social relations, and this context is essential for conducting accurate, ethical analysis.
7. **Make labor visible.** The work of data science, like all work in the world, is the work of many hands. Data feminism makes this labor visible so that it can be recognized and valued.

## RESTORATIVE/TRANSFORMATIVE DATA SCIENCE EXAMPLES

What does a restorative/transformational data science project look like? Across my research into femicide data activism practices and my participation in various communities of practice, I have made some basic and preliminary observations about characteristics that many restorative/transformational data science projects share.

A restorative/transformational data science project

- has a theory of power and a theory of change;
- intentionally uses a data epistemology that focuses on liberatory goals (e.g., healing, liberation, emancipation, sovereignty, refusal, self-determination);
- can be done with minimal computing resources and basic data literacy skills;
- can be done with small, medium, or big data;
- can be done *without* an advanced degree from a fancy institution;
- often (but not always) produces or uses counterdata;
- engages in ethical, long-term relations of care with the communities most impacted;
- engages in pluralistic and culturally appropriate conceptions of rigor and truth; and
- holds the responsibility of caring for the data and the people and the stories and the relations assembled in the database.

All the grassroots femicide monitoring projects I have discussed in this book are examples of restorative/transformational data science (indeed, they are the motivation for theorizing the concept in the first place). Examples from other domains abound. Ahead are projects across domains ranging from housing to health to the environment to civic engagement. Each one does not necessarily share all of the characteristics listed previously, but each has something to offer in terms of epistemology, thematic focus, community engagement, research methods, or outputs. I offer these as models to serve as inspiration for your own projects.

## HOUSING

### Anti-Eviction Mapping Project

<https://antievictionmap.com/>

This activist-academic project defines evictions more broadly than the legal definitions and uses their discordant data to advocate for a more structural framing of the root causes of eviction in specific areas. They acquire official eviction data from court records and other eviction data from surveys and collaborations with housing clinics. They have produced maps of evictions, landlord monitoring tools, murals,

oral histories, and a book called *Counterpoints: A San Francisco Bay Area Atlas of Displacement & Resistance*.

### **“The Secret Bias Hidden in Mortgage-Approval Algorithms”**

<https://themarkup.org/denied/2021/08/25/the-secret-bias-hidden-in-mortgage-approval-algorithms>

The result of a year-long investigation by the Markup, this data journalism story demonstrates systemic racial bias in home mortgage loan approvals in the United States. The journalists used publicly available official data triangulated with academic research studies, but there are still key missing data hindering a comprehensive analysis—notably credit scores—which the mortgage industry has successfully lobbied to keep secret.

## **LAND USE**

### **The Detroit Geographic Expedition and Institute**

<https://medium.com/nightingale/gwendolyn-warren-and-the-detroit-geographic-expedition-and-institute-df9ee10e6ad2>

The Detroit Geographic Expedition and Institute (DGEI) was a collaboration between Black young adults in Detroit led by Gwendolyn Warren and white academic geographers that lasted from 1968 to 1971. The group worked together to produce data about aspects of the urban environment related to children and education, and produced numerous maps from that work, including the widely circulated map *Where Commuters Run Over Black Children on the Pointes-Downtown Track*. The DGEI collected many other types of counterdata and published them in reports with analysis and recommendations.

### **Woorani territorial mapping / Mapeo Territorial Woorani**

<https://waoresist.amazonfrontlines.org/explore/>

The Woorani people’s lands are in the upper part of the Amazon River in Ecuador. Threatened by oil companies, the Woorani began a process of mapping their knowledge of and relation to the land in order to “defend our way of life and protect our future from threats like oil exploitation, mining impacts and invasions.”<sup>9</sup> They use a participatory process that involves training different groups to use GPS devices, produces large paper maps that can be discussed and annotated in community, and finally leads to publication of the resulting maps on the open-source platform Mapeo. Some maps are public, others are not, and “everything included on the map is the cultural property of the Woorani.”

**“Land Grab Universities”**

<https://www.landgrabu.org/>

*High Country News* produced an original report and unique database documenting how land grant universities across the United States were funded with expropriated Indigenous land via the 1862 Morrill Act. Data sources included land patent records, congressional documents, historical bulletins, historical maps, and more. Many data needed manual entry, and the database took months to assemble. It is publicly available for download and research.

**HEALTH****Mapping Police Violence**

<https://mappingpoliceviolence.org/>

Run by advocacy organization Campaign Zero, this project has tracked and mapped fatal police violence—and its systemic racial injustice—in the United States since 2013. Similar to femicide data activists, the project relies on media reports as a primary source and triangulates those with official data and other counterdata sources. The database is open and publicly available.

**“Lost Mothers”**

<https://www.propublica.org/article/lost-mothers-maternal-health-died-childbirth-pregnancy>

Data on maternal mortality in the United States have been characterized as “an unreliable mess” by *Scientific American*.<sup>10</sup> In 2016, ProPublica set out to identify every single mother or parent who died from pregnancy-related causes in the United States (estimated to be between seven hundred and nine hundred people). They used social media, crowdfunding sites where funds had been set up for families left behind, public records, and obituaries. The journalists discuss how their crowdsourcing approach ended up overrepresenting the stories of dominant groups—the white, educated women who were more likely to respond to their call for stories—and underrepresented Black women in particular. Adriana Gallardo wrote about these methods in “How We Collected Nearly 5,000 Stories of Maternal Harm.” This is an important reminder of the limitations of counterdata tactics and the ways in which they may reproduce the matrix of domination.

**COVID Black**

<https://covidblack.org/>

An organization founded during the COVID-19 pandemic out of a national campaign for people to call and demand that US state and federal agencies collect

and publish racial data. Founded by historian and digital humanities scholar Kim Gallon, COVID Black not only gathers and publishes Black health data, but also does trainings and produces data stories, visualizations, and other public interpretations to combat racial health inequities (and also to push back against the relentless, racialized deficit narratives depicted by the mainstream media). They state on their website, “Data is more than facts and statistics. Black health data represents life.”

### **The Qanuippitaa? National Inuit Health Survey**

<https://nationalinuithealthsurvey.ca>

Qanuippitaa? National Inuit Health Survey (QNIHS) is an ongoing longitudinal survey of the health and well-being of the Inuit—the Indigenous peoples of the Arctic. It is Inuit-owned and Inuit-determined, and works in partnership with four major Inuit land claims organizations to ensure that the survey, data collection, data analysis, and research outputs are owned by Inuit people and informed by Inuit knowledge, values, and worldview.

## **ENVIRONMENT**

### **The Global Atlas of Environmental Justice**

<https://ejatlas.org/>

Initiated in 2012 by a team of researchers at the Universitat Autònoma de Barcelona, in Spain, the Environmental Justice Atlas undertakes systematic collection of global ecological conflicts in partnership with activists, civil society organizations, and social movements. They source conflicts from media reports, crowdsourcing, and local partnerships with impacted communities. The project has developed its own typology of ecological conflict and publishes its data openly.

### **Environmental Data & Governance Initiative**

<https://envirodatagov.org/>

Environmental Data & Governance Initiative (EDGI) is a research collaborative that was formed in 2016 in the United States as the Trump administration threatened to take open federal environmental datasets offline (a notable example of intentionally producing missing data). The group started organizing “data rescues” where they downloaded and archived federal datasets on university and civil society servers. Now EDGI monitors US government action (and inaction) related to environmental information, organizes campaigns, and educates communities who want to use data to lobby the government. They have also led the

development of environmental data justice principles (see the data epistemology section).

### **Land and Environmental Defenders Campaign**

<https://www.globalwitness.org/en/campaigns/environmental-activists/>

An advocacy campaign, open dataset, and series of reports undertaken by Global Witness to record the unjust deaths of people—largely Indigenous people—killed while defending their land and environments. Similar to femicide data activists, Global Witness sources cases from social media, news reports, and trusted local partners and networks. They have a strict information verification methodology to determine if a case should be included in their database. They outline that their numbers are “only a partial picture” because of the difficulty of obtaining comprehensive information about these cases.

## **TRANSPORTATION**

### **Assaults on ride-hail drivers**

<https://themarkup.org/newsletter/hello-world/tracking-tens-of-thousands-of-assaults-on-ride-hail-drivers>

Since 2021, the Markup has been monitoring assaults and carjackings on Uber and Lyft drivers, some of which end in fatalities. Reporter Dara Kerr sourced hundreds of these assaults from phone calls, police reports, public information requests, and local news articles. The data are published in an open, searchable database. Affected individuals and labor organizations are using these numbers to call for more gig worker safety protocols and protections.

### **The DTP Map / Карта ДТП**

<https://dtp-stat.ru>

An activist map and ongoing monitoring effort in Russia that combines official government data on traffic crashes with weather, streets, participant types, and more, as a way of instigating civic engagement and social change to reduce accidents. Here the activists are not producing their own counterdata but rather reframing official data as a way of building a public sense of urgency around traffic fatalities. Yet the activists have mixed feelings about the veracity of the official data, as described in a case study by Dmitry Muravyov, “Doubt To Be Certain: Epistemological Ambiguity of Data in the Case of Grassroots Mapping of Traffic Accidents in Russia.”



## ARTS AND CULTURE

### **National Monument Audit**

<https://monumentlab.com/audit>

A study by the arts-based nonprofit Monument Lab found that the monument landscape in the United States is overwhelmingly white and male and elevates themes of war and conquest. There is no authority in the United States that keeps records on monuments, so to undertake the national audit, they assembled records from dozens of federal, state, local, tribal, and institutional sources, many of which have different criteria and definitions of “monument.” They held participatory analysis sessions to develop their analytical categories and themes. The data can be explored in their interface on the project’s website.

### **Whose Heritage? Public Symbols of the Confederacy**

<https://www.splcenter.org/whose-heritage>

Since 2015, the Southern Poverty Law Center (SPLC) has maintained a database and map of Confederate-related monuments and place names. They periodically reaudit the list to monitor removals, relocations and renamings. For example, following George Floyd’s murder, the SPLC found that almost one hundred Confederate symbols were removed, relocated, or renamed. They publish reports as well as a community action guide that allows users of the map to take direct action by providing instructions on how to build a campaign to remove monuments and/or rename streets.

### **Book Censorship Database**

[https://www.everylibraryinstitute.org/book\\_censorship\\_database\\_magnusson](https://www.everylibraryinstitute.org/book_censorship_database_magnusson)

A project by Dr. Tasslyn Magnusson in partnership with EveryLibrary Institute and EveryLibrary to monitor book bans and book challenges across the United States since 2021. The open spreadsheet is organized by school districts, books challenged/banned in school districts, public libraries, and books banned/challenged in school libraries. Data are sourced from news reports, online forums, and social media.

### **Abortion Onscreen Database**

<https://www.ansirh.org/research/abortion/pop-culture>

The Advancing New Standards in Reproductive Health (ANSIRH) organization compiles and publishes an open database of all film and television depictions available to viewers in the United States that discuss abortion, from 2016 to the present. ANSIRH publishes annual reports on media portrayals of abortion, and the database has been used in a range of media studies.

## CIVIC ENGAGEMENT

### Feminindex

<https://winguweb.org/en/caso-impacto/feminindex/>

A civic media project by Ecofeminita and Wingu that documented and visualized where political candidates in Argentina stood on gender and LGBTQ+ issues, including reproductive rights, femicide, care work, and trans rights. The first version was released in 2017, with subsequent versions in 2019 and 2021. Data on politicians' views were collected through media reports, candidates' public statements, and policy documents and through surveys administered by the organization.

### First Nations Information Governance Centre

<https://fnigc.ca>

First Nations Information Governance Centre (FNIGC) is a nonprofit organization leading the establishment of Indigenous data sovereignty for all members of the Assembly of First Nations in Canada. They undertake a variety of ongoing First Nations population surveys, run trainings and capacity-building sessions for tribal partners, and develop data governance strategy aligned with the goals of sovereignty and self-determination.

## HOW TO USE THIS TOOLKIT

<b>RESOLVING</b>  Developing a theory of change	<b>RESEARCHING</b>  Finding + verifying information	<b>RECORDING</b>  Information extraction + classification	<b>REFUSING + USING</b>  Where data go, who uses them
---	---	---	---

Although there are many ways the toolkit could be used, here are three possibilities:

### 1. Strategic project planning and visioning

Use the toolkit at the beginning of a project to undertake planning sessions, establish a shared vision for the project, who it is serving, and how to sustain it through the different stages of work. In this model, the project team would take the time to go through the questions and activities in this restorative/transformational data science toolkit, as individuals and as a group, align around their answers at each stage, and incorporate them into their plan and their vision (and their budget!).

## 2. Equity pauses and recalibrations

I learned about the idea of *equity pauses* from Jenn Roberts, who runs VersedEd and the Colored Girls Liberation Lab. This is the idea of regularly stepping back from the intense day-to-day work, say, of researching and recording counterdata, and pausing to evaluate your process and whether you are meeting your equity goals. In this model, teams would take a short period of time to engage with the questions at one stage of work, discuss their answers, and surface shifts and recalibrations to make in data practices to better meet their goals. These equity pauses could happen at regular, scheduled intervals—for example, at one meeting a month.

## 3. Ethics crisis moments

There may be moments in a restorative/transformative data science project that provoke an equity pause that the team did not foresee. A community may come forward and express that they have been harmed. An individual's information may have been made public in a traumatizing way. You or your team may have included a story or a case or some information without permission. These are moments where a more profound recalibration—of data practices and of relationships—becomes necessary. This toolkit could aid in that recalibration by providing a structured set of ethical questions and activities for the team to use to draw out their analysis of what happened, how to redress it in the short term, and how to prevent such harm in the longer term.

## GETTING STARTED

To get started, I suggest teams first do *Start-up Activity 1: What is your data epistemology?* to reflect on who they are, who they serve, and who they are accountable to with their data work. This will help you understand which emerging alternative data epistemologies may match with your team's backgrounds, relations, goals and values. It will also help you begin to reflect on and refine your theory of change for your work. *Start-up Activity 2: Map the information ecosystem* guides your group to create a map of the information ecosystem for your topic of interest. Understanding the information ecosystem will be invaluable as you begin to think about how your group can mitigate biases, address missing data, build coalitions, and use data and information for healing and liberation. The information ecosystem map that you make during this activity is referenced in a number of later activities, so it is handy to keep around as a guide for your project.

### START-UP ACTIVITY 1: WHAT IS YOUR DATA EPISTEMOLOGY?

**Time required:** 60–75 minutes

**Materials:** Pens and paper/sticky notes

**Preactivity homework:** Review the list of data epistemologies provided in this toolkit (or other data epistemologies you may be considering).

**Activity:** Choosing an appropriate data epistemology involves locating yourself and your team and your organization/institution in relation to the topic. These questions can serve as a starting point and are designed to be answered as a group. For each question, take five minutes to quietly freewrite answers, and then ten to fifteen minutes to share responses with the group.

1. Who are you (individually and as a team and as an organization) in relation to the topic? Do you bring lived experience of the topic? How and why were you brought to the topic?
2. Who are you producing data for? Which communities or publics do you serve, or aspire to serve, by doing this work?
3. Who are you accountable to? Which communities or publics should have a direct say in influencing the course of the project?

**Final discussion (15 minutes):** Review your responses in relation to the list of data epistemologies in this toolkit. Which one or ones are best aligned with your team's responses?

### START-UP ACTIVITY 2: MAP THE INFORMATION ECOSYSTEM

**Time required:** 2½–3 hours

**Participants:** 5+. Can be done with fewer than five people but it will take more time. If you are able, try to recruit participants with lived experience, legal experience, data experience, and movement experience.

**Materials:** Computers with Internet access, multicolored sticky notes, six large posterboards or papers

**Preactivity homework:** Organizers and team members should determine the geographic scale of interest and the time period of interest for the project and write it into a concise mission statement in this form: "We are mapping the information ecosystem for TOPIC in PLACE from START TIME to END TIME."

**Activity setup:**

1. Write out your mission statement (see preactivity homework) on a blackboard, whiteboard, or big paper posted for everyone to see.

2. Choose one color of sticky notes that will represent “missing data” and one color that will represent “bias” and communicate those to the participants.

**Activity:**

1. 15 min.: Introductions—participants go around the room and say their name, pronouns, and one personal or professional reason they are in the room today.
2. 10 minutes: Organizers read the mission statement and outline the purpose for gathering today. Divide participants into groups and give each group a large surface to work on (posterboard or wall) and sticky notes.
3. 60 minutes: Group work.
  - **Group 1—Legal inventory:** Place a sticky note on your surface for each law relevant to your topic. Note the year it was passed and whether it is a municipal/state/federal/tribal law on the sticky note. Include laws that relate to (1) definitions of the phenomenon, (2) government or official monitoring of the phenomenon, and (3) public disclosure of information about it. As you examine each law, place a “bias” sticky note if you see a source of bias.
  - **Group 2—Official data producers inventory:** Place a sticky note on your surface for each agency or group that produces official information about the topic of interest. Around it, place another sticky note for each relevant dataset that the agency produces. Note on the sticky note whether that dataset is open or closed, aggregated or disaggregated. If the dataset is open, download it and examine it. As you examine each agency and dataset, place a “bias” sticky note if you see a source of bias and a “missing data” sticky note if you see missing datasets, rows, features, or variables.
  - **Group 3—Counterdata producers inventory:** Place a sticky note on your surface for each agency or group that produces counterdata or activist data about the topic of interest. Note on the sticky note what sector they are from—e.g., activism, journalism, nonprofit, academia, government, etc. Around it, place another sticky note for each dataset that that group produces. Note on the sticky note whether that dataset is open or closed, aggregated or disaggregated. If the dataset is open, download it and examine it. As you examine each group and dataset, place a “bias” sticky note if you see a source of bias and a “missing data” sticky note if you see missing datasets, rows, features, or variables.
  - **Group 4—Data users inventory:** Place a sticky note on your surface for each agency or group that *uses* data (official data or counterdata) about the topic of interest. Note on the sticky note what sector they are from—e.g., government, activism, journalism, nonprofit, academia, etc. As you examine each data user, place a “bias” sticky note if you see a source of bias.
  - **Group 5—Larger landscape inventory:** Place a sticky note on your surface for other actors in this information ecosystem that are influencing policy, advocacy, and public conversation. They may not produce or use data, but they are individuals, organizations, government agencies, and/or social movements that are doing agenda-setting

on the topic. As you examine each individual or group, place a “bias” sticky note if you see a source of bias.

**Stretch break! (5 minutes)**

**Discussion & reflection (60 minutes):**

- 30 minutes: Each group shares back their results (5 min. per group)
- 20 minutes: Large group power analysis. Facilitators move participants through the following questions:
  - What data remain missing from this mapping? What data should exist (according to your group or to other advocacy groups) but do not exist in either official or counterdata efforts? Why?
  - What are the biases in this information ecosystem? What are their root causes? Can the biases and inequalities be mitigated informatically?
- 10 minutes: Close out and designate participants who can help document the work.

**Post-activity documentation:** Consider creating a large visual map synthesizing groups’ work that can be posted in your team’s space (digital or physical space).

Once you have done the two start-up activities, your team is ready to pick and choose from the rest of this toolkit to see which of the activities and discussions might be relevant for your project. Throughout this book, I have described the different workflow stages of a restorative/transformational data science project: resolving, researching, recording, and refusing and using data. These stages are derived from our team’s interviews with grassroots data activists working to challenge feminicide, predominantly from Latin America. These stages form a four-stage process model, and I offer that model here in the hopes that it may be useful to other practitioners working on restorative/transformational data science projects.

Different ethical concerns and questions arise at each stage of work in a restorative/transformational data science project. For example, during the resolving stage of a project, data practitioners are developing their analysis of the problem, their theory of change for how and why counting and data analysis might be useful, and their data epistemology. Here it is important to think about your own positionality in relation to the phenomenon, who the beneficiaries of a project may be, and how to work collectively, in networks. In contrast, during the research stage of a project, it is important to reflect on the systemic biases in the information ecosystem, creative ways to source information, and how your project will handle missing data about minorities and subgroups. The rest of the toolkit is structured around activities to do and questions to ask during these four stages of a restorative/transformational data science project.

As you will see ahead, each discussion or activity is mapped to the data feminism principle that it aligns with. This is my attempt to demonstrate how one’s data epistemology can translate into concrete matters of discussion and action for people working with data. Following the publication of *Data Feminism*, many people have asked Lauren and me for practical guidance on how to use the data feminism principles—to move them from general guidelines into something applicable in specific contexts. This mapping is an attempt to do that, as well as a way of inviting scholars and activists to do this kind of mapping work for other data epistemologies.

<p><b>RESOLVING</b></p> <p>Developing a theory of change</p>	<p><b>RESEARCHING</b></p> <p>Finding + verifying information</p>	<p><b>RECORDING</b></p> <p>Information extraction + classification</p>	<p><b>REFUSING + USING</b></p> <p>Where data go, who uses them</p>
--	--	--	--

*Resolving* is the stage of a restorative/transformative data science project in which an individual or group seeks to address a problem of structural inequality and determines how and why counting and registering data will be an effective method to do so.

<p><b>Examine power</b></p>	<p>Look at the information ecosystem map you generated in Start-up Activity 2. How can you <i>join</i> rather than <i>initiate</i> your own effort? [ACTIVITY]</p> <p>Whose job is it to count? Whose job are you doing? How will you remind them that it is their job? [DISCUSSION]</p> <p>Are you from the community you are counting? Are you working within institutional structures that incentivize knowledge extraction (academia, journalism, nonprofit) or profit (industry)? Are you the right person or group to be doing this work? Try to have the courage to <i>not</i> do the project. [DISCUSSION]</p> <p>Which ideas—from theorists, activists, and/or communities—will you draw from for your structural analysis of power? How does your counting work help to name and frame the phenomenon as a structural problem? [DISCUSSION]</p>
<p><b>Challenge power</b></p>	<p>Who are you counting for? [DISCUSSION]</p> <p>What is your theory of change? How and why do you think measuring and monitoring will challenge power? [DISCUSSION]</p> <p>What is your data epistemology? See Start-up Activity 1 in this toolkit if you don’t know yet. Once you do know, create a collaborative document with resources and guidance about your chosen data epistemology to help get new team members up to speed with its foundational ideas and methods. [ACTIVITY]</p>

What are ways that counting may harm the people and communities you want to serve (e.g., by making them visible to institutions that want to target them)? [DISCUSSION]

Is this a one-time study or an ongoing observatory? How does that match your available resources (people and money)? How does that match who you want to serve or influence? [DISCUSSION]

---

**Consider context**

How can your project center the lived experience of those who have been impacted by the issue? Without exploitation, extraction, or tokenism? [DISCUSSION]

What is specific to the geography or community that you are counting? What differences do you need to highlight? Do you need to develop new names, frames, concepts, and/or categories for that context? [DISCUSSION]

---

**Embrace Pluralism**

How can you avoid hoarding, whether data or credit? How can you count in community—leveraging collectives and networks of solidarity? How can you build partnerships and work in networks instead of trying to do everything on your own? [DISCUSSION]

---

**Elevate emotion and embodiment**

What kind of emotional labor is involved in this work? How will you care for yourself and your team as you measure injustice? [DISCUSSION]

---

<b>RESOLVING</b>  Developing a theory of change	<b>RESEARCHING</b>  Finding + verifying information	<b>RECORDING</b>  Information extraction + classification	<b>REFUSING + USING</b>  Where data go, who uses them
---	---	---	---

*Researching is the stage of a restorative/transformational data science project in which an individual or group seeks and finds data observations and related information to add to their database. This can include sourcing existing datasets, discovery and detection of new observations, triangulation of information across sources, and ongoing research to add information to existing observations.*

**Examine power**

Discuss the following questions with your team before beginning research:

- What harm does collecting counterdata potentially incur for the groups that you want to influence or serve?
- Whose trauma does it make visible and what is your (individual/team's/institution's) relationship to that trauma?
- Could bad actors use your data to target minoritized groups?
- Do the data perpetuate a deficit narrative about minoritized groups—painting them as in need of saving by dominant groups?

[DISCUSSION]



	Return to the biases surfaced on your information ecosystem map (Start-up Activity 2). What are their root causes? How can your work name and challenge those biases? [DISCUSSION]
Consider context	<p>Return to missing data surfaced on your information ecosystem map. Discuss with your team members the following question: “why don’t these data exist?” [DISCUSSION]</p> <p>Augment your information ecosystem map by placing a sticky note to denote creative ways that you can navigate, mitigate, and triangulate missing data. Don’t forget to consider mass media, hyperlocal media, social media, private chat groups, relationships, partnerships, friendships, and crowdsourcing. [ACTIVITY]</p> <p>What groups, especially those at the intersection of multiple forms of domination, will still be missing, underreported, erased, or neglected by your methods of counterdata research? How can you address those limitations or, at the very least, acknowledge them? [DISCUSSION]</p>
Embrace pluralism	How can you cultivate human networks of counterdata research predicated on ethical, authentic, nonextractive, and enduring relations? These might be relations with individuals, social movements, coalitions, journalists, nonprofits, or service organizations from your information ecosystem map. [DISCUSSION]
Elevate emotion and embodiment	How emotionally challenging is the research? Consider what it might be like for survivors or people with first-hand experience. How will you handle self-care and team care for secondary trauma? What does a trauma-informed approach to the production of this data look like? [DISCUSSION]
Make labor visible	How can you make the labor of researching counterdata visible and for whom? Are there strategic reasons for hiding the labor of your counterdata research? Are there ways to acknowledge labor and care internally, even when it may be strategic to conceal them externally? [DISCUSSION]

RESOLVING	RESEARCHING	RECORDING	REFUSING + USING
Developing a theory of change	Finding + verifying information	Information extraction + classification	Where data go, who uses them

*Recording is the stage of a restorative/transformative data science project that involves extracting unstructured data from various sources into structured datasets (text documents, spreadsheets, and/or databases); classifying cases according to diverse typologies; and managing data—including ethics, access, and governance of the database.*

Examine power	<p>What does consent look like for the type of data you are collecting? How will you obtain consent and also support individuals and groups who decide to withdraw their consent? [DISCUSSION]</p> <p>Take a look at the legal landscape inventory from your information ecosystem map in Start-up Activity 2. How are legal protections falling short? Do laws adequately capture and classify the scope of the phenomena? [DISCUSSION]</p>
Challenge power	<p>How can you count and classify in order to exceed and/or challenge those standards? How might you demand that the phenomenon be conceptualized and measured differently? [DISCUSSION]</p> <p>Look at your information ecosystem map. What important variables and categories are missing from existing data? How can you incorporate those into your recording work? [DISCUSSION]</p>
Embrace pluralism	<p>How can your counterdata project operate as a <i>megaphone</i> for amplifying the voices, power, knowledge and agency of the people closest to the harms that you are trying to challenge? [DISCUSSION]</p> <p>How will you engage multiple and diverse stakeholders in the development of your data variables and categories? How will you participate in building community—the essential, ongoing social and technical infrastructure that can sustain this work? [DISCUSSION]</p> <p>Who else is recording counterdata about the issue? How can you scale your impact by <i>harmonizing</i> with them—that is, sharing definitions, categories, dialogue, and recording tips, and even potentially pooling data for greater impact? Look at your information ecosystem map for groups to start with. [DISCUSSION]</p>
Rethink binaries and hierarchies	<p>Whose experiences are sidelined, erased, or marginalized by your schema and categories? Whose experiences will be sidelined because there will be quantitatively fewer of them in the dataset and/or because of known biases in the data sources? How do you bring these experiences back in? [DISCUSSION]</p> <p>If necessary to the project, how can you collect identity categories such as race, gender, and ethnicity without naturalizing and essentializing them? Which categories might you avoid collecting because to collect them would be to do harm (e.g., trying to record someone’s race or gender from a photo of them)? [DISCUSSION]</p>
Elevate emotion and embodiment	<p>How will you care for and respect your data? How will you develop your team’s intimate knowledge of and relationship with the data? [DISCUSSION]</p> <p>How do your columns and categories communicate certain narratives about the issue and the people involved? (For example, is a woman always named as a “victim,” defining her life and her agency by a single event?) How can you push back on that essentializing tendency? [DISCUSSION]</p>

How is your database a memorial to structural trauma—a *cultural countermemory*? Whose lives and whose pain is represented therein? How are you accountable to them and how are you in relationship with them? [DISCUSSION]

If your database is a memorial, how does this shift your thinking about ethics and access to the database? [DISCUSSION]

---

**Make labor visible**

What is the minimal computing infrastructure you need for your team to do the work easily, safely, and reliably? How can you balance minimal computing and easy-to-use tools with data security and redundancy? [DISCUSSION]

---

<b>RESOLVING</b>  Developing a theory of change	<b>RESEARCHING</b>  Finding + verifying information	<b>RECORDING</b>  Information extraction + classification	<b>REFUSING + USING</b>  Where data go, who uses them
---	---	---	---

*Refusing and using data is the stage of a restorative/transformative data science project in which individuals and groups circulate data in order to push specific actors toward thinking, feeling, or acting differently. The goals of these data actions and circulations may include to repair, to remember, to reframe, to reform, and/or to revolt.*

**Challenge power**

Freewrite or free-draw about *refusal* and the issue area that you are working on. What are you refusing? Who is refusing? What is the affirmative, generative vision forged from your refusal?

(For example, for femicide data activists, the affirmative vision is a world that has eradicated gender-related violence and its causal forces of oppression: cisheteropatriarchy, settler colonialism, white supremacy, racial capitalism, and more.) [ACTIVITY]

Data may refuse the status quo in a variety of ways. Here are some questions to start thinking about the ways your project may use data for refusal:

- **Repair:** How can data support the provision of direct services, support, and healing to the people most impacted by the issue?
- **Remember:** How can data support the holding of collective space around a structural problem? How can data reshape collective memory and communal interpretation of an issue?
- **Reframe:** How can data's authority be leveraged to amplify grassroots voices and analysis in media and culture? To use narrative change to do mass public consciousness-raising?

- **Reform:** How can data be leveraged to design new or reform existing institutions, laws, policies, and official practices? How can data’s authority be leveraged to get a seat at the table for impacted communities?
- **Revolt:** How can data be used in support of massive mobilizations in public spaces or online? How can data become spectacular and evocative, physically or digitally, to claim space, to claim time, to demand public attention, and to enact dissent? [DISCUSSION]

Embrace pluralism	Review your information ecosystem map and make a list of the different groups or audiences that you want to move to action and brainstorm multiple forms of data communication tailored for each audience. (For example, if you want to move policy makers to action, one form might be a report with data visualizations. Another form might be oral testimony from impacted communities. Another form might be a visual slideshow with photos, quotes, and statistics. Another form might be a protest outside their offices.) Each form is an opportunity to involve different groups in the communication and the circulation of data. [ACTIVITY]
Elevate emotion and embodiment	<p>When is it politically advantageous to communicate data neutrally and minimally, as if from an omniscient observer? When is it more appropriate to center emotion and embodiment in data communication? [DISCUSSION]</p> <p>What emotional impact will your data artifacts have when circulated publicly? Is the impact different for different groups (say, survivors or impacted communities)? How will you care for the impact on those most affected and give them avenues for healing and action? [DISCUSSION]</p>
Consider context	How can you recontextualize your data points and recuperate them from their abstraction into rows and columns? How can acts of data communication and circulation connect each data point back into the fullness of the lifeworlds from which it emerged? [DISCUSSION]

### A LIVING TOOLKIT

This is a first step toward a toolkit for restorative/transformative data science. I welcome feedback and dialogue. Given that many of the examples I provide are from the United States, I would especially love to learn about restorative and transformative data science projects outside of the US context. Please post all comments, questions, and critiques online to the evolving open toolkit located at <https://mitpressonpubpub.mitpress.mit.edu/pub/restorative-data-toolkit>.

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

# Counting Feminicide

## Data Feminism in Action

By: Catherine D'Ignazio

### Citation:

*Counting Feminicide: Data Feminism in Action*

By: Catherine D'Ignazio

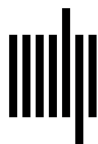
DOI: 10.7551/mitpress/14671.001.0001

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